Deficits in Theory of Mind and Executive Function

as Risk Factors for Conduct Problems

from Middle Childhood to Early Adolescence –

a Longitudinal Perspective

Dissertation

zur Erlangung des akademischen Grades Doktor der Philosophie (Dr. phil.)

Eingereicht bei der Humanwissenschaftlichen Fakultät

der Universität Potsdam

von

Anna Katharina Holl (Dipl-Psych.)

Potsdam, Oktober 2018

Gutachterinnen:

1. Gutachten: Prof. Dr. Birgit Elsner
2. Gutachten: Prof. Dr. Annette Klein

Disputation:

Am 17.10.2018 an der Universität Potsdam.

Published online at the
Institutional Repository of the University of Potsdam:
https://doi.org/10.25932/publishup-45991
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-459915

“Ever tried. Ever failed. No matter. Try again. Fail again. Fail better.”

Samuel Beckett

Acknowledgements

I want to thank my supervisors, Birgit Elsner and Barbara Krahé, for the opportunity to write this thesis, and for their guidance and support along the way.

Very special thanks go to my colleagues from the PIER Study: Finja Gallit, Svenja Hoffmann, Fabian Kirsch, Nele Lensing, Sebastian Löweke, Susanne Meiser, Johannes Meixner, and Greta Warner. Thank you for being such a great PIER group!

Further, I also want to thank my colleagues from the developmental department for their emotional and scientific support: Maurits Adam, Matt Hilton, Viola Meckelmann, Christiane Patzwald, Bernd Schellhas, Maria Schönebeck, and Markus Studtmann.

Last, but not least, I want to express my deep gratitude to all my wonderful friends and family, who supported me these past years while I was writing this thesis, and particularly Jonas, who always believed in me and gave me the courage to continue.
Without you this would not have been possible!

Finally, I have to thank Tinka, Leela, and Lilly, for their patience and understanding that I had to sit in front of the laptop for so many hours ...

Table of Contents

1. Introduction

In the last few decades, increasing rates of psychiatric disorders in children have been reported (e.g., Sourander et al., 2008), and prevalence rates of conduct problems more than doubled from 1974 to 1986 for 15-year-olds and from 1974 to 1999 for 16-year-olds (Collishaw, Maughan, Goodman, & Pickles, 2004). Conduct problems include disobeying rules, aggression, property destruction, stealing, and bullying (Bevilacqua, Hale, Barker, & Viner, 2017). Prevalence rates of conduct problems increased equally for both genders, for all family types, and across all social classes in Western industrialized countries. Whether this rise in prevalence rates is indeed a consequence of an increase in the population prevalence of this serious mental health problem is under debate. Some argue that rising prevalence rates could also be a consequence of increased clinical recognition, changing diagnostic criteria or investigation of different age groups (Collishaw, 2012). This is supported by a recent study that identified declines in mean conduct problem scores in 7-year-olds in Great Britain (Sellers, Maughan, Pickles, Thapar, & Collishaw, 2015). However, long-term consequences of serious conduct problems in childhood, such as repeated offences, irregular employment, violence in relationships, and as a consequence social exclusion (Farrington, 1995), still call for preventive action, not only to alter the quality of life for those affected, but also to reduce the financial burden on the public.

The cost for society caused by female conduct disorder is – contrary to earlier speculation – about the same as for male conduct disorder (Scott, Knapp, Henderson, & Maughan, 2001), although the prevalence rates are usually lower for females. For instance, in a German representative sample of 948 children, 15.5% of children, 14.8% of girls, and 16.1% of boys between 7 and 10 years old were reported by their parents to have conduct disorder symptoms above the cut-off in the German version of the CBCL (Ravens-Sieberer et al., 2008). If the criterion of present impairment is considered, 8.7% of children, 6.7% of girls, and 10.5% of boys were reported to exhibit symptoms above the cut-off.

The long-term consequences and prevalence rates emphasize the necessity of studying early risk factors for conduct problems in children, because a sufficient understanding of the etiology is a prerequisite for maximizing the effect of interventions that prevent conduct problems (Beauchaine & Hinshaw, 2015). A lot of knowledge has already been gathered about environmental risk factors, such as neighborhoods with low socioeconomic status (e.g., Dodge, Pettit, & Bates, 1994), that play an important role in the development of conduct problems. On the other hand, evidence of intrapersonal factors that influence the development of conduct problems is not yet conclusive. Since Crick and Dodge (1994) formulated their social information-processing (SIP) model, research has focused on the study of cognitive factors that are associated with aggressive behavior in children, particularly in the form of biases. As aggressive behavior is a significant part of conduct problems (see diagnostic criteria in 2.1.1), the SIP model builds a theoretical foundation for the present thesis. The model proposes that children who show more aggressive behavior than their peers deviate in their SIP from non-aggressive children. For example, they are proposed to attribute hostile intentions to their mates (e.g., de Castro, Veerman, Koops, Bosch, & Monshouwer, 2002), even when the latter do not have such intentions. Consequently, after going through a process of six different steps in their information processing, the children conclude that behaving aggressively is the best solution in a particular situation. These processes are explained in detail in section 2.1.6.

The reason why it is still necessary to study intrapersonal cognitive factors that affect the development of conduct problems is that since then new advances on the side of cognitive development research have been made, in particular in the study of theory of mind (see 2.4), and executive function (see 2.5). Theory of mind (ToM) involves making inferences about the perspectives of others in a social situation (Premack & Woodruff, 1978), and is one of the most important functions of our social brain (Frith & Frith, 2010). Executive function (EF) describes a set of general purposeful control mechanisms that regulate human cognition and behavior and are linked to the prefrontal cortex of the brain (Miyake et al., 2000). They make human life as we experience it possible, because they enable us to concentrate and pay attention to social interactions, when going on automatic would be ill-advised or impossible (Diamond, 2013). Perhaps because these two cognitive constructs were not on the research agendas when Crick and Dodge (1994) formulated their model, they were not literally included in it. However, these two cognitive domains are promising as playing an important role in SIP, because research has confirmed that children with deficits in ToM and EF show more aggressive behavior than their non-aggressive peers (see sections 2.4.6 and 2.5.4); however, other research has found no or a positive relation between theory of mind or executive function and aggressive behavior and conduct problems. Consequently, it is not yet clear whether or not deficits in theory of mind and executive function are risk factors for conduct problems and aggressive behavior.

In addition, despite being interrelated throughout development (see 2.3.6), deficits in theory of mind and executive function have mostly been studied separately as risk factors for aggressive behavior and conduct problems. Particularly, in the developmental period of middle childhood and in the transition to adolescence, where different developmental changes on cognitive, emotional, and social levels take place (Collins, 1984; Steinberg & Morris, 2001), they have mostly been examined in separate samples or in separate analyses. Consequently, the present thesis focuses on deficits in both theory of mind and executive function as risk factors for conduct problems in middle childhood and early adolescence. This is done by adopting a longitudinal perspective and research design, as the developmental taxonomy of conduct disorder suggests (Moffitt, 1993; see 2.1.2), as well as transactional perspectives on externalizing behavior (Beauchaine, Shader & Hinshaw, 2015) and the SIP model (Crick & Dodge, 1994). This allows the study of reciprocal processes throughout development (see 2.1.5 and 2.1.6).

The research questions for the present thesis are explained in detail in section 2.4, and the three studies that have been conducted are presented in sections 3, 4, and 5, and some additional analyses in paragraph 6. Afterwards, the results are discussed in section 7, and finally a conclusion is drawn in paragraph 8.

2. Theoretical Background

2.1. Conduct problems

2.1.1. Definition of conduct problems

Conduct problem behaviors in middle childhood and adolescence include aggression towards people and animals, destruction of property, deceitfulness or theft, and serious violation of rules (e.g., Lahey et al., 2000). These are also criteria for the diagnosis of conduct disorder, a psychiatric syndrome that involves recurrent and persistent conduct problems in children and adolescents (Blair, Leibenluft, & Pine, 2014), but which is less common than conduct problems with a prevalence rate between 2% and 16% in community samples (Loeber, Burke, Lahey, Winters, & Zera, 2000). Conduct disorder is defined as a repetitive and persistent pattern of behavior that violates the rights of others, and is not in accordance with major age-appropriate societal norms and rules (Kimonis & Frick, 2010). However, the clinical diagnosis of conduct disorder can only be determined via a clinical assessment that delivers sufficient diagnostic information. For instance, only by such an assessment can the criterion of clinically significant impairment in social, academic, or occupational function be evaluated, which has to be fulfilled in order to diagnose conduct disorder (American Psychiatric Association [APA], 2013). Furthermore, at least three symptoms must have been present in the past 12 months, with at least one in the past 6 months, to diagnose conduct disorder.

In the present thesis, the focus lies on conduct problems; however, due to great overlap in the literature on conduct problems and conduct disorder, literature on the latter will also be reviewed. Furthermore, conduct disorder and antisocial or dissocial behavior are often used interchangeably in the literature. However, while conduct disorder refers to the psychiatric disorder, antisocial or dissocial behavior usually refers to behavior that is hostile or harmful to organized society, or behavior that deviates sharply from the social norm (Merriam-Webster, 2018). Influential literature on antisocial behavior will also be considered, such as the developmental taxonomy of antisocial behavior by Moffitt (1993). In this context, the term 'antisocial behavior' will be used, otherwise the term 'conduct problems' is used for the above-described behaviors and 'conduct disorder' for the psychiatric diagnosis.

Furthermore, some of the studies on conduct problem behaviors have examined not only conduct disorder, but also oppositional defiant disorder. Symptoms for oppositional defiant disorder are organized into three groups, namely angry/irritable mood, argumentative/defiant behavior, and vindictiveness (APA, 2013). Together with conduct disorder, attention-deficit/hyperactivity disorder, substance use disorder, antisocial personality disorder, and disruptive mood dysregulation disorder, oppositional defiant disorders are often classified as disruptive behavior disorder or externalizing spectrum disorders (Drabick, Steinberg, & Hampton, 2015). In earlier versions of the Diagnostic and Statistical Manual of Mental Disorders (DSM), such as in the DSM-IV, oppositional defiant disorder could not be diagnosed in the presence of conduct disorder (Drabick et al., 2015). This criterion was omitted in the most recent version of the DSM, the DSM-5, because children who show symptoms of oppositional defiant disorder do not necessarily have conduct disorder or progress to it, although most adolescents who have conduct disorder have previously met diagnostic criteria for oppositional defiant disorder (Beauchaine & McNulty, 2013). One reason for this could be that children and adolescents with conduct problems seem to follow distinct developmental trajectories, which will be explained in the next paragraph.

2.1.2. Developmental trajectories of conduct problems

A widely known developmental taxonomy conceptualized by Moffitt (1993) still plays a fundamental role for theory and research. Originally, it differentiated between two types of antisocial behavior in youths. These two types, which are also included in the diagnostic criteria and etiological theory of conduct disorder (DSM V) (e.g., Burke, Loeber, Lahey, &

Rathouz, 2005), were proposed to have different etiologies and to be characterized by different outcomes in adulthood. Children from the childhood-onset group are theorized to engage in antisocial behavior before the age of 10 and follow a life-course-persistent path of antisocial behavior. The early signs of conduct disorder that these children show involve aggressive behavioral tendencies, impulsivity, and failure to comply with requests, all of which are also symptoms of attention deficit-hyperactivity disorder and oppositional defiant disorder (Blair et al., 2014). Typically, children, mostly boys, in the early-onset group follow a trajectory from attention deficit-hyperactivity disorder in early school years to oppositional-defiant disorder and subsequent conduct disorder with the beginning of adolescence (e.g., Burke et al., 2005). However, most children with oppositional defiant disorder do not develop conduct disorder, and oppositional defiant disorder is also predictive of anxiety and depression in boys. The best predictor of conduct disorder in boys one year later are previous conduct disorder symptoms, but oppositional defiant disorder symptoms are a second significant predictor (Burke et al., 2005).

By contrast, as defined in the developmental taxonomy of antisocial behavior by Moffitt (1993), adolescence-limited offenders exhibit antisocial behavior only during adolescence and are the largest group of antisocial children. This is reflected in a steep increase in the prevalence of antisocial behavior in adolescence. However, the exact prevalence rate depends on the source of data on antisocial behavior; for instance, self-reports yield higher rates than official records on court convictions or police arrests (Moffitt, Caspi, Rutter, & Silva, 2001). In the Dunedin longitudinal study, a large epidemiological study on antisocial and delinquent behavior situated in Dunedin, New Zealand, the propositions of the developmental taxonomy were examined. One result was that childhood predictors differed between delinquents on the life-course-persistent path and delinquents on the adolescence-limited path (Moffitt & Caspi, 2001). The former exhibited more risk factors, like inadequate parenting, neurocognitive risks, and temperament and behavioral problems in childhood. The reason for the latter, namely, why some adolescents show antisocial behavior that is limited to adolescence, is seen in the so-called 'maturity gap' (Moffitt, 1993). The maturity gap describes the discrepancy in modern societies between the age of biological maturity and social maturity, that is, the age at which adult behaviors like drinking alcohol and driving are legal for adolescents. Further, the influence of deviant peers increases with adolescence. They serve as role models that demonstrate how to cope with the discomfort of the maturity gap by means of delinquent behavior (e.g., Elliot & Menard, 1996). In adolescence, a shift is proposed in which deviant peers receive increasing influence and higher status in the peer group, which serves as an additional reinforcement for delinquent behavior (Moffitt, 1993). These two distinct developmental trajectories are also recognized in the DSM V, in which an additional specification is necessary: If individuals show at least one symptom prior to age 10, the criterion for childhood-onset is fulfilled, and if not, the criterion for adolescent-onset type is met, or, when not enough information is available, the onset is unspecified (APA, 2013).

In addition, Moffitt and colleagues already identified a third group of antisocial children whose behavior problems start in childhood but do not continue through to adolescence, and which was therefore named 'recovery group' or childhood-limited pathway (Moffitt, Caspi, Dickson, Silva, & Stanton, 1996). Since then, many population-based studies have confirmed this third trajectory of antisocial behavior, and have actually found that 50%-70% of children with childhood-onset antisocial behavior have outgrown their difficulties by adolescence (e.g., Barker & Maughan, 2009; Nagin & Tremblay, 1999; Odgers et al., 2008).

Further, a recent review focuses on differences between children and adolescents with conduct disorder who show remorse, empathy, and concern about school performance in contrast to children and adolescents who do not show these and instead display callous-unemotional traits (Blair et al., 2014). The latter occur in fewer than half of those with conduct disorder and are correlated with poorer treatment response and prognosis (Blair et al., 2014). Adult outcomes of the different conduct problems trajectories differ depending on the specific trajectory group in accordance with Moffitt's taxonomy (1993): Individuals who have shown conduct behavioral problems from early on and follow a persistent course of conduct problems display the highest risk of poor outcomes in early adulthood or later in comparison to a group without significant conduct problems (e.g., Bevilacqua, Hale, Barker, & Viner, 2017). Examples of poor adult outcomes include serious mental health problems like depression, substance abuse, self-reported aggression, an official record of antisocial behavior, and poor general health. One particularly problematic outcome is antisocial personality disorder, a serious pathological disorder which is associated with high levels of persistent aggression and rule-violating behavior and broadly interferes with adaptive social functioning. It develops in slightly less than 50% of people who were previously diagnosed with conduct disorder and is associated with callous-unemotional traits (Blair et al., 2014). Individuals of the adolescent-onset trajectory group exhibit intermediate risk across most problematic adult outcomes, and individuals of the childhood-limited trajectory had the lowest odds ratios compared to the two former groups (Bevilacqua et al., 2017). However, adults of the last trajectory, who only exhibited conduct problems in childhood, are still at a higher risk in comparison to those in the low trajectory on self-reported aggression and poor educational outcomes (e.g., Bevilacqua et al., 2017). As a consequence, it was proposed that not the age of onset of conduct problems alone is a decisive predictor of adult outcomes, but rather that the course of conduct problems across childhood and adolescence is most predictive of later outcomes. In sum, an implication is that the developmental periods of middle childhood and early adolescence are important target periods for the study of conduct disorder and problems, because with the beginning of adolescence an increase in conduct symptoms can be observed (Moffitt et al., 2001). In addition, it would be beneficial to use longitudinal designs that allow for the study of developmental trajectories of conduct problems.

2.1.3. Gender differences in conduct problems

The developmental taxonomy of antisocial behavior proposes that the different trajectories of antisocial behavior, that is, childhood-onset and adolescence-onset delinquency, apply to both males and females (Moffitt, 1993). However, as males exhibit more neurodevelopmental vulnerabilities and risk factors in early childhood, in childhood-onset delinquency the male-to-female ratio in the prevalence of conduct disorder is usually 10:1 (Moffitt & Caspi, 2001), while in adolescence-onset antisocial behavior, the sex ratio is 1.5:1. This is explained by the maturity gap, the main cause for adolescence-onset delinquency, which should apply equally to males and females in adolescence. In line with this proposition, in the Dunedin sample each girl's delinquency was associated with the timing of her puberty (Moffitt et al., 2001). Altogether, evidence from the Dunedin longitudinal study confirms that the majority of females who exhibit delinquency are in the adolescence-onset trajectory group, whose delinquent behavior has started after the age of 10 and after the beginning of puberty (Moffitt & Caspi, 2001). By contrast, evidence from another longitudinal study that focused on girls' conduct disorder found that the most common age of onset for girls was before the age of 10 (Keenan, Wroblewski, Hipwell, & Stouthamer-loeber, 2010). One reason for these contrasting results could be that this study did not extend far enough into adolescence, because it only included girls up to the age of 15, and at this wave they had fewer observations than before; therefore, the estimates could have been biased toward the younger ages. Consequently, based on other studies on conduct problems and disorder, Keenan et al. (2010) argued that the most common pathway towards conduct disorder for girls is via an exacerbation of symptoms between childhood and adolescence.

When it comes to gender differences in conduct symptoms and etiological factors, research is scarce, and in general great overlap is expected (Vloet, Großheinrich, Konrad, Freitag, & Herpertz-Dahlmann, 2014). However, some gender differences in neurobiological correlates of conduct problems have been found, for instance in autonomic correlates, which appeared as followed: Boys who show aggressive behavior with conduct problems exhibit reduced autonomic functioning, whereas girls with aggressive behavior demonstrate greater electrodermal responding than controls (Beauchaine, Hong, & Marsh, 2008). Further, some evidence points to differences in environmental risk factors between girls and boys. For example, in a longitudinal study, higher rates in externalizing behavior at age 9 were predicted by the child's male gender and lower levels of maternal sensitivity, particularly in boys (Miner & Clarke-Stewart, 2008). In addition, regarding callous-unemotional traits, environmental risk factors seem to have a greater impact than genetic risks for girls in comparison to boys (Fontaine, Rijsdijk, McCrory, & Viding, 2010). Nevertheless, epigenetic factors could also play an important role for girls' conduct problems (Vloet et al., 2014). In sum, more research is needed, particularly on conduct problems in girls and on gender differences in conduct behavioral symptoms and etiological factors.

2.1.4. Aggressive behavior

2.1.4.1. Forms of aggressive behavior

Aggressive behavior forms one cluster of symptoms of conduct problems (APA, 2013), and aggressive behavior can also be differentiated into distinct forms and functions. Aggressive behavior is defined as "any form of behavior directed toward the goal of harming or injuring another live being who is motivated to avoid such treatment" (Baron & Richardson, 1994, p. 7). Considering the forms of aggressive behavior, harming others is possible through physical means, which is known as physical aggression (e.g., Crick, Casas, & Ku, 1999). However, harming others can also be done through more relational acts, that is, by damaging friendships or feelings of inclusion in a peer group, such as by spreading rumors, which is called relational aggression (Crick & Grotpeter, 1995). The latter is sometimes referred to as indirect aggression or social aggression with an emphasis on distinct behaviors; however, all three constructs essentially refer to the same kind of aggression (Archer & Coyne, 2005). Therefore, in the following only the expression 'relational aggression' is used, even when cited research and authors use an alternative term.

The developmental theory of aggression by Bjorkqvist and colleagues proposes that the use of aggression is normative, and which type of aggression is used depends on distinct phases in socio-cognitive development (Björkqvist, Österman, & Kaukiainen, 1992). Empirical evidence on the developmental course of the forms of aggression has confirmed that the majority of children, as expected, decrease their use of physical aggression after the age of 2, simultaneous with cognitive developments (Côté, Vaillancourt, Barker, Nagin, & Tremblay, 2007). At the same time, about two-thirds of children exhibit low levels of relational aggression, and about one-third follow high rising trajectories from age 2 to 10 (Côté et al., 2007; Vaillancourt, Miller, Fagbemi, Côté, & Tremblay, 2007). Considering combined levels of both forms, most children (62.1%) exhibit decreasing levels of physical aggression and low levels of relational aggression. Further, a significant proportion shows moderately decreasing physical aggression and increasing relational aggression (14.2%), and 13.5% exhibit high levels of both forms of aggression (Côté et al., 2007).

In a longitudinal study on the development of relational aggression in middle childhood, gender differences in the developmental course between the ages of 8 and 11 were found (Spieker et al., 2012). Boys had lower initial levels and decreasing rates of relational aggression, whereas girls exhibited higher initial levels, but no significant change. Regarding gender differences in physical aggression, boys typically engage in physically aggressive behavior more often than girls (e.g., Lansford et al., 2012). Different theoretical explanations for this gender difference include evolutionary, hormonal, and social role accounts (see Krahé, 2013, for a review).

By contrast, evidence of gender differences in relational aggression is mixed. Several studies found that girls show more relationally aggressive behavior than boys (e.g., Ostrov, Murray-Close, Godleski, & Hart, 2013). This can be explained by girls' tendency to have more dyadic relationships than boys, which seems to fuel relational aggression (Murray-Close, Ostrov, & Crick, 2007). However, other studies reveal no gender differences (e.g., Lansford et al., 2012), or only trivial gender differences with girls showing more relational aggression (see Card, Stucky, Sawalani, & Little, 2008, for a meta-analysis), and one study even finds more relationally aggressive behavior in boys (Henington, Hughes, Cavell, & Thompson, 1998). Altogether, particularly regarding relational aggression, further research is needed on the developmental course of aggression during middle childhood including both forms of aggression, and an investigation of gender differences.

2.1.4.2. Functions of aggressive behavior

Different functions of aggression are distinguished in relation to the motivation that drives a person to act aggressively. Reactive aggression, which is also called "hot-blooded" or defensive aggression, refers to angry responses to others' behavior that is perceived as threatening, and is often associated with anger (e.g., Card & Little, 2006). By contrast, proactive aggression, also called instrumental or "cold-blooded" aggression, refers to goal-oriented, deliberately calculated behavior that is motivated by external rewards (e.g., Dodge, Lochman, Harnish, Bates, & Pettit, 1997). Proactive aggression can be explained in light of Bandura's (1973, 1986; cited after Card & Little, 2006) social-cognitive learning theory, which explains aggression as a product of high self-efficacy, favorable outcome expectations, and positive valuing of outcomes that are reached through aggressive means. Proactive aggression has been linked to maladaptive outcomes, including delinquency and antisocial personality disorder, and conceptually aligns with callous-unemotional traits (e.g., Vitaro, Brendgen, & Tremblay, 2002; Vitaro, Brendgen, & Barker, 2006). On the contrary, reactive aggression is a direct response to frustration as incorporated in the frustration-aggression hypothesis (Berkowitz, 1989). Reactive aggression has been linked to differential outcomes and behavior, particularly to aggression against familiar persons, for instance, dating violence, in comparison to proactive aggression (Brendgen et al., 2001; Hubbard, McAuliffe, Morrow, & Romano, 2010). Despite these different explanations and psychosocial outcomes, reactive and proactive aggression are correlated with each other, though the direction and strength of the relation varies considerably from -.10 to .89 (Polman, de Castro, Koops, van Boxtel, & Merk, 2007). A recent view conceptualizes the two subtypes as continuous dimensions that exist in each child to varying degrees (Hubbard et al., 2010).

One of the few longitudinal studies on the developmental course from 5th to 9th grade determined that both reactive and proactive aggression peak in 6th grade and decline afterwards (Fite, Colder, Lochman, & Wells, 2008). However, in this study, only 126 children screened as highly aggressive were included. Another longitudinal study on 1,571 Swiss children identified six dual trajectories of reactive and proactive aggression over the course of six years, which were as follows: the largest group (56%) with low, stable levels of both aggression types, a second group with moderate levels of both subtypes of aggression with slightly higher levels of aggressive behavior during preadolescence (8.5%), a third group with moderate initial levels of reactive aggression, which decreased from middle childhood to early adolescence, but low levels of proactive aggression (19.7%), a fourth group with high initial levels of both functions of aggression but decreases across time (6.4%), a fifth group with high levels of . In sum, the largest group of children showed low levels of both reactive and proactive aggression; however, different subgroups show variation and change throughout middle childhood and early adolescence. Consequently, the developmental periods of middle childhood and early adolescence seem particularly worthwhile when studying the functions of aggression.

Research on gender differences in reactive and proactive aggression is very scarce, and very little is known about females who engage in reactive and/ or proactive aggression, because of the few studies that exist some focus solely on males (e.g., Ellis, Weiss, & Lochman, 2009). This emphasizes the need to include females in the study of the functions of aggression.

2.1.5. Etiology of conduct problems

Research from recent decades suggests at least four different subtypes of conduct disorder that all show distinct developmental trajectories and etiological processes, among them affective, biological, cognitive, familial, and environmental risk factors (Kimonis & Frick, 2010). The first and second are the childhood-limited and adolescence-onset subtypes of conduct disorder described above. Further, in the childhood-onset conduct disorder a distinction is made between children who show low levels of emotional reactivity to cues of others' distress, which is called 'callous-unemotional traits' or 'limited prosocial emotions' (this is the term used in the DSM V), and children without such traits. This separates the third and fourth types of conduct disorder. Altogether, multiple developmental pathways can lead to conduct disorder, and different etiological factors may be responsible for distinct forms of conduct problems.

In the developmental taxonomy of conduct disorder, it has already been proposed that individuals in the childhood-onset persistent trajectory group display a higher number of risk factors, such as neurocognitive difficulties, which include deficits in EF and poorly controlled behavior, although those in the adolescence-limited group should not show these risk factors (Moffitt, 1993). These propositions were confirmed in the Dunedin longitudinal study, in which children who had exhibited serious signs of conduct disorder before the age of 10 proved to show more risk factors on the side of the child, such as neuropsychological deficits, in combination with more environmental risks like inadequate parenting (Moffitt & Caspi, 2001). However, recent research suggests that differences between childhood-onset and adolescence-onset antisocial behavior are more quantitative than qualitative (Fairchild, Van Goozen, Calder, & Goodyer, 2013). The adolescence-onset type may also display a neurodevelopmental disorder with more negative future outcomes than healthy individuals, similar to the childhood-onset type. Thereby, the quality of children's early environment is supposed to moderate between individual vulnerabilities and age-of-onset of antisocial behavior (Fairchild et al., 2013).

Further, data from genetic epidemiological studies have shown moderate to high heritability for clinical aspects of conduct problems (Blair et al., 2014), for instance, aggression (Rhee & Waldman, 2002), and others have found that rule-breaking behavior is less heritable than aggression (Edelbrock, Rende, Plomin, & Thompson, 1995). Recent research also links specific neurocognitive dysfunctions with clinical aspects of conduct disorder, such as deficient empathic responding (Blair et al., 2014). Deficient empathy is related to an impaired mechanism that should inhibit violence (e.g., Richell et al., 2003). Empathy is critical for socialization and, similarly, taking another's perspective, a facet of ToM, is important for intentional moral conduct (e.g., Hoffman, 1987). Children who show conduct disorder, particularly with higher levels of callous-unemotional traits, display an impaired ability to recognize distress in peers (e.g., Marsh & Blair, 2008). Deficient empathy has been related to reduced amygdala functioning (e.g., Marsh et al., 2008), which is why this reasoning is also called the 'amygdala theory of violent behavior' (Sharp, 2008).

Transactional theories of developmental psychopathology propose that different sets of risk factors may interact in specific ways with each other and still lead to similar psychopathological problems (i.e. aggressive behavior and conduct problems), which is called 'equifinality' (Gatzke-Kopp, Greenberg, Fortunato, & Coccia, 2012). By contrast, the opposite is also observed, namely that children with similar clinical profiles often have distinct longitudinal outcomes in conduct disorder (Blair et al., 2014), known as ‘multifinality’ (Cicchetti & Rogosch, 1996). A recent ‘ontogenic processes model of externalizing psychopathology’ views psychopathological developments as a consequence of complex longitudinal transactions between interdependent vulnerabilities on the individual level (e.g., genetic, epigenetic, neural) and equally interdependent contextual risk factors (e.g., coercive parenting, deviant peer group affiliations, etc.) (Beauchaine, Shader, & Hinshaw, 2015). From this transactional perspective, a given predisposition in a child, such as a cognitive deficit, may elicit certain responses from others (i.e., the peers), which then feeds back into the child’s development and behavior (Kimonis & Frick, 2010). This idea of reciprocal processes was already included in the influential model of SIP (Crick & Dodge, 1994) to explain aggressive behavior in children, which will be explained in the next paragraph, and is an important implication for the present thesis.

2.1.6. The SIP model

Since Crick and Dodge (1994) formulated their theory on SIP in children, different studies have covered the relation between components of SIP and antisocial behavior in children. The authors hypothesized then that children’s behavior is a function of their individual processing of social cues, and that specific processing styles could lead to maladjustment, particularly to aggressive behavior in children (Crick & Dodge, 1994). According to their theory, children’s cognitive SIP can be depicted in six different steps (see Figure 1): First, *encoding* of both internal and external cues (Step 1), to which children attend selectively, takes place. In the following *interpretation* (Step 2), different independent processes may occur, including a mental representation of the situational cues, a causal analysis of the events in the present situation, inferences about the perspectives of others in the situation and their intentions, and other interpretative processes, such as evaluations of past performances. Then, children select a goal or desired outcome for the situation, which is called *clarification of goals* (Step 3). Goals include focused arousal states, like keeping calm or staying out of trouble, which means that arousal regulation is included in goal selection. In the next step, *response access or construction*, children access possible responses to the present situation from memory or construct new behaviors (Step 4). The generated responses are evaluated and the most positively evaluated response is chosen in the *response decision* to be enacted (Step 5). Finally, the chosen response is *enacted behaviorally* (Step 6). Naturally, social interaction and mental processing do not stop at this point, but a new cycle of processing may start here.

Further, Crick and Dodge (1994) assumed that social information processes are influenced by past social experiences. They proposed that memories are integrated into a 'database,' also called *latent mental structures*. In the database, memories are transformed into social schemas, scripts, working models of relationships, cognitive heuristics, and social knowledge (Dodge, 2006). These are accessed in each of the proposed steps and consequently influence the processing and behavior in a present situation. At the same time, the memories of past social encounters, including aggressive behavior, are stored in the database, and influence future cognitive SIP. These developmental cascades have been investigated only in one longitudinal study up to now. In this study, it was determined that deficits in SIP and aggressive behavior in children are linked through peer rejection in children from kindergarten through grade 3 (Lansford, Malone, Dodge, Pettit, & Bates, 2010). Deficits in SIP are also related to concurrent externalizing problem behaviors in children and adolescents in particular (e.g., Lansford et al., 2006). A typical deficit or bias in SIP is to attribute hostile intentions to other children in provocative, ambiguous situations, which is common in aggressive children (e.g., Crick & Dodge, 1994; 1996) (see paragraph 2.2.6.).

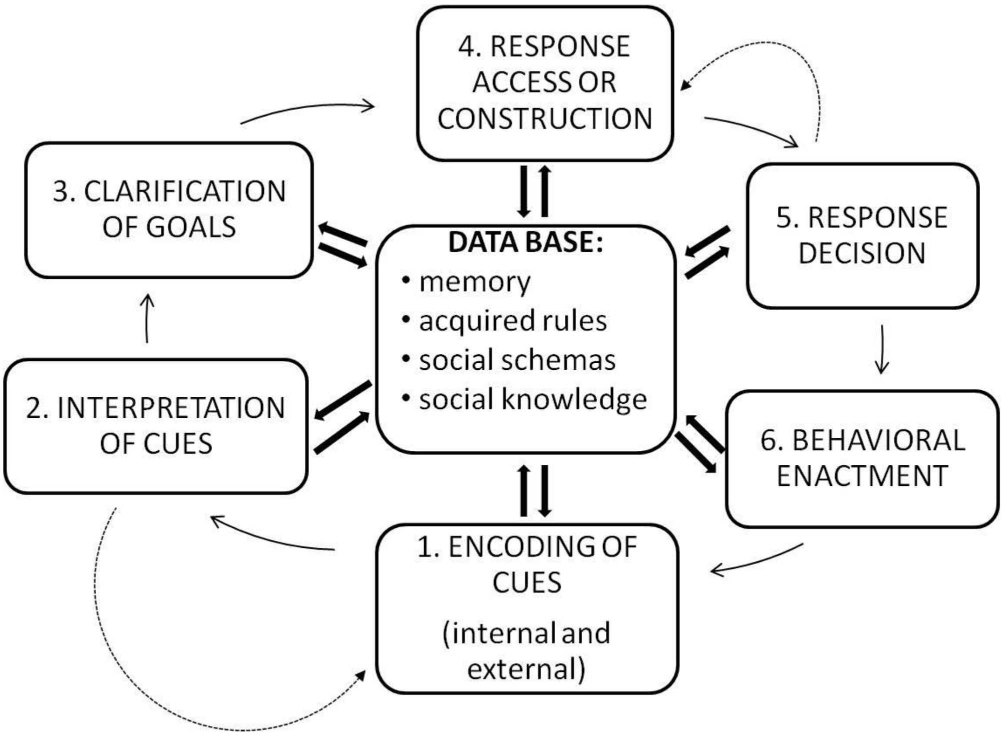


Figure 1. SIP model freely depicted following Crick & Dodge (1994)

2.1.7. Emotions in the SIP model

It has been recognized that – in addition to cognitions – emotions play an important part in SIP as well (Lemerise & Arsenio, 2000). As a consequence, emotional processes that encompass emotionality, emotion regulation, and moods or background emotions have been integrated in the SIP model of Crick and Dodge (1994). Since emotionality as a temperament style or other emotional processes are a predisposition that a child brings into a social interaction, these emotional processes are proposed to mediate between social-cognitive processes and behavior in a given situation and the social knowledge that is stored in the database. Further, additional emotional processes are integrated in every step of SIP, for instance empathic responsiveness in Step 1, or the affective nature of the relationship with a peer in Steps 2, 3, and 5 (Lemerise & Arsenio, 2000).

One emotion that is of particular importance in the study of aggression is anger. It has been found that children who are prone to anger are more likely to engage in aggressive behavior (e.g., Hubbard et al., 2002; Olson, Sameroff, Kerr, Lopez, & Wellman, 2005). In middle childhood, anger shares genetic covariance with conduct problems (Deater-Deckard et al., 2007). Research on irritability, defined as increased proneness to anger as a psychological trait (Leibenluft & Stoddard, 2013), found that irritability predicted disruptive behavior symptoms, including aggressive behavior, in early childhood (Wakschlag et al., 2015). Reactively aggressive behavior is usually seen as the direct consequence of frustration (see 2.3.2.), as explained by the aggression-frustration hypothesis (Berkowitz, 1989). Also, irritability is conceptualized as an aberrant response to frustration or threat (Leibenluft, 2011). Additional research revealed that anger arousal and aggression-related motor impulses, that is, an action tendency, are closely related (Berkowitz & Harmon-Jones, 2004a). Further, an approach motivation that is linked to feelings of anger fosters reactively aggressive behavior (Berkowitz, 2012). In light of the adapted SIP model by Lemerise and Arsenio (2000), anger seems to mediate between social-cognitive processes, such as hostile attributional bias, and reactively aggressive behavior in children (e.g., de Castro, Merk, Koops, Veerman, & Bosch, 2005; Dodge et al., 2015).

2.2. Theory of mind (ToM)

2.2.1. What is ToM?

Since Premack and Woodruff (1978) discussed whether chimpanzees have a theory of mind, the study of theory of mind has become one of the most popular topics in cognitive development in recent decades (e.g., Flavell, 2004; Miller, 2009). “Theory of mind concerns an understanding of mental state, what we know or believe about thoughts, desires, emotions, and other psychological entities both in ourselves and in others” (Miller, 2009, p.1). Having a ToM means understanding that what we think, believe, desire, and feel may be different from the mental states of other people, and also from reality. ToM – which is also called ‘mentalizing’ – enables individuals to infer others’ intentions and plans, and to “predict what others are going to do on the basis of their desires and beliefs" (Frith & Frith, 2010). The cognitive processes that are attributed to ToM allow for deliberate communication, teaching, and cooperation (Frith & Frith, 2012), and consequently are basic processes for social interaction.

2.2.2. Differentiating cognitive from affective ToM

Recently, a distinction has been made between cognitive and affective theory of mind (ToM). According to Shamay-Tsoory and Aharon-Peretz (2007), cognitive ToM is defined as inferences made regarding others' beliefs, and affective ToM as inferences made regarding others' emotions.

Shamay-Tsoory and colleagues proposed that, although cognitive and affective ToM are partly dependent on each other, they are dissociable and depend on distinct neuroanatomical substrates (Shamay-Tsoory, Tibi-Elhanany, & Aharon-Peretz, 2006; Shamay-Tsoory and Aharon-Peretz, 2007). In their research on brain lesions, they found that these substrates also overlap, as both cognitive and affective ToM depend on intact prefrontal cortex functioning. However, they determined that cognitive ToM is affected by widespread damage in prefrontal regions, particularly in the ventromedial and dorsolateral cortex, whereas affective ToM is impaired from circumscribed ventromedial prefrontal damage. Further research revealed that criminal offenders with psychopathic features and individuals with lesions in the orbitofrontal cortex both showed impairment in affective ToM (Shamay-Tsoory, Harari, Aharon-Peretz, & Levkovitz, 2010). They proposed that cognitive ToM is a prerequisite for affective ToM, and that affective ToM is linked to cognitive empathy. In their view, the concepts of affective ToM and cognitive empathy are quite similar, because they both relate to the understanding of others' emotions. Contrary to that, an emotional empathic ability theorized to be related to emotional contagion is located in the inferior frontal gyrus (e.g., Shamay-Tsoory, Aharon-Peretz, & Perry, 2009).

2.2.3. Development and assessment of ToM in early childhood

ToM is not part of the cognitive repertoire from birth, but is rather a developmental achievement (Miller, 2009). The first wave of studies has concentrated on what is termed 'false belief.' This was first conceptualized by Wimmer and Perner (1983), who tested in their famous 'Maxi-Task' whether young children aged 3 to 9 understood that a protagonist had a false belief about the location of an object. While the children saw that mother moved the chocolate to a different cupboard, Maxi, the protagonist in the story, did not know about the transfer of the chocolate and consequently, would look in the wrong location for the object. However, when asked where Maxi would look for his chocolate, the 3-to-4-year-old children all answered that Maxi would look in the right location, which is of course the wrong answer, because Maxi should not know about the new location. The reason for their mistake is that they fail to represent the false belief of the protagonist correctly. 3-year-olds typically fail a false belief task, which is usually of similar content to the one described, while some 4-year-olds exhibit success and 5-year-olds typically succeed (for a meta-analysis, see: Wellman, Cross, & Watson, 2001). The mastering of first-order false belief is considered one of the milestones of preschool cognitive development (Miller, 2009).

Before this age, some precursors of ToM have been found, for instance infants' social attention, which predicts later social cognition (Wellman, Phillips, Dunphy-Lelii, & LaLonde, 2004). Some researchers argue that infants already have a ToM (e.g., Onishi & Baillargeon, 2005), in contrast to others who argue that infants' good performance on tasks of paying attention to human faces and motion instead reflects an implicit understanding of behavior in infants (e.g., Ruffman, 2014). This is still an issue of much debate; however, the focus of the present thesis lies on ToM development in middle childhood and early adolescence.

2.2.4. Development and assessment of ToM in middle childhood and adolescence

It is widely accepted now that there is developmental life in ToM after first-order false belief (Miller, 2009). One the one hand, there are studies on the so-called second-order false belief, which are appropriate for children between about 5 and 6 or 7 years old, and as such are situated during the transition to middle childhood, depending on the applied task. On the other hand, there are other, advanced tasks that capture individual differences across distinct aspects of ToM (e.g., cognitions, emotions, perspectives, desires, intentions) in older children and adolescents. In recent years, a few more studies have focused on advanced tasks in ToM research (Hughes & Devine, 2015).

The first kind of studies investigating second-order false belief directly builds on first-order false belief tasks. They also typically use an analogous story paradigm; however, a second character is necessary. For instance, the belief of Maxi's mother, who does not know that Maxi has returned and sees the chocolate in the other location, could be added to the story. When the children are asked after the story what Maxi's mother believes about Maxi's belief, this is already second-order reasoning. This approach was also first developed by Perner and Wimmer (1985), and many other studies have adopted the vignettes they introduced. Children around the age of 6 or 7 years old typically master this task under optimal conditions, when inference of second-order beliefs is prompted (Miller, 2009). Other stories that are quite frequently applied include those developed by Sullivan, Zaitchik, and Tager-Flusberg (1994). Moreover, they are simpler and include a procedural variation, and most children already succeed in them at the age of 5 (for a review, see Miller, 2009).

Other tasks to assess higher-order reasoning use different approaches. Some of them were initially created for the study of autism and also use a story paradigm, like the Strange Stories by Happé (1994). This task involves the detection of sarcasm, bluff, irony, and double bluff, which are not well understood until the age of 8. Another task that compares typically developing children aged 7 to 11 years old with children with high-functioning autism or

Asperger Syndrome assesses ToM through recognition of faux pas, also using a story paradigm (Baron-Cohen, O'Riordan, Stone, Jones, & Plaisted, 1999). A faux pas might be committed when a speaker says something without considering whether it might offend the listener. However, some assessments use other media, such as the Silent Film task (Devine & Hughes, 2013), which is supposed to be a film-based analogue of the Strange Stories task. Another famous task is the Eyes task that uses a series of photos depicting the eye region of the face (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). A variation on this is the Faces task, in which video clips are shown that depict the torsos and faces of actors who display different advanced emotional states (Golan, Baron-Cohen, & Hill, 2006). In both – the Eyes and the Faces task – the task is to identify what the person depicted is thinking or feeling. Such assessments that demand a judgment of emotions, like in the Faces task, can be attributed to assessing affective ToM. By contrast, tasks that demand judgments of thoughts, beliefs or desires – like the traditional false belief task – assess cognitive ToM.

Recent research from brain imaging (e.g., Blakemore & Choudhury, 2006; Choudhury, Blakemore, & Charman, 2006) and behavioral studies (e.g., Apperly, Warren, Andrews, Grant, & Todd, 2011; Bosacki & Astington, 1999; Devine & Hughes, 2013; Dumontheil, Apperly, & Blakemore, 2010; Vetter, Leipold, Kliegel, Phillips, & Altgassen, 2013) supports the idea that ToM, in particular affective ToM (e.g., Sebastian et al., 2012; Vetter, Weigelt, Döhnel, Smolka, & Kliegel, 2014), continues to develop from middle childhood and adolescence through to young adulthood. However, most of the cited research concentrates on adolescence, and only very few studies investigate ToM development in middle childhood and in children who would be expected to pass simple first- and second-order false belief tasks (Apperly et al., 2011). Because of ceiling effects, the usage of tasks originally developed for younger children is restricted. Consequently, only a limited number of tasks are available to provide sufficient variability and determine ToM development in middle childhood and adolescence. Therefore, new tasks, like the above-mentioned Faces task, need to be developed further, and their application has to be tested.

Another possibility for circumventing the problem of restricted variance in ToM tasks for older children is the use of reaction times. For example, Apperly et al. (2011) used reaction times, in addition to error rates, in their belief-desire reasoning task, which allowed the detection of developmental continuity. They compared three groups of children and found that the 10- to 11-year-olds showed fewer errors than the 8- to 9-year-olds, who in turn showed fewer errors than the 6- to 7-year-olds. Further, they determined that young children, older children, and adults find it more difficult to reason about false beliefs and negative desires than true beliefs and positive desires.

Altogether, in the last few years different studies have confirmed ongoing development in ToM in middle childhood and adolescence, however most of it has concentrated on adolescence. Consequently, more research is needed to investigate ToM development in middle childhood and early adolescence (preadolescence), and for this aim different assessment paradigms are needed in comparison to early childhood.

2.2.5. Gender differences in ToM

In early childhood, gender differences have remained rarely investigated until now. One of the few studies that did examine them found no differences between female and male preschoolers' ToM (Hughes, Ensor, & Marks, 2011); however, one study found female superiority in two out of three ToM tasks in 4- to 6-year olds (Villanueva Badenes, Clemente Estevan, & Garcia Bacete, 2000). Research on gender differences in older children is also limited. Some studies have only investigated female participants (Dumontheil et al., 2010; Vetter et al., 2014) due to basic gender differences in brain development. In the few studies that investigated gender differences, girls generally scored higher than boys on ToM tasks (in 7-, 9-, and 11-year-olds: Baron-Cohen et al., 1999; in 10- to 13-year-olds: Bosacki &

Astington, 1999; in 6- to 8-year-olds: Calero, Salles, Semelman, & Sigman, 2013; in 8- to 13-year olds: Devine & Hughes, 2013; in 12- to 13-year-olds: Ibanez et al., 2013). In a study that additionally investigated mentalizing speed, girls also outperformed boys (Keulers, Evers, Stiers, & Jolles, 2010). By contrast, in one study no differences between female and male preadolescents were found in the "Reading the emotions of the eyes" task (in 9- to 13-year-olds: Sharp, 2008).

Different theoretical reasons are considered for female superiority on ToM tasks. Baron-Cohen (2002) claims that the brains of males and females are distinct in two dimensions: 'empathizing' and 'systemizing.' Typically, females should be better at empathizing, which includes identifying others' emotions and thoughts and predicting behavior, all of which are abilities needed for ToM. By contrast, males should be better at systemizing, which includes the application of rules and analyzing systems. As a consequence, in general females should outperform males on tasks that involve empathy and ToM. Further research has revealed that prenatal biology in the form of fetal testosterone may also influence abilities, such as empathy and ToM, mediated by androgen effects in the brain (Chapman et al., 2006). However, this would not explain the lack of gender differences in preschoolers' ToM.

Devine and Hughes (2013) speculated that female superiority on ToM tasks may only manifest in more challenging tasks like the ones that are applied to older children, such as their Silent Films task. One possible explanation for this developmental difference comes from Maccoby's (e.g., 1990) theory of gender differences in children's play behavior. Girls typically prefer activities that promote verbal communication, like pretend play, whereas boys typically prefer activities that promote spatial abilities, like playing Cowboys and Indians (e.g., Pasterski, Golombok, & Hines, 2011). These preferences are proposed to contribute to later differences in cognitive performance.

In sum, from the limited evidence it seems that in middle childhood and early adolescence girls outperform boys on most advanced ToM tasks, but further research on this topic is needed. Theoretical explanations for this gender difference include biological and social factors, but more empirical support from longitudinal studies is needed.

2.2.6. Deficits in ToM as a risk factor for conduct problems

As mentioned before (see paragraph 2.1.6), in the SIP model by Crick and Dodge (1994), as well as in the adapted model by Lemerise & Arsenio (2000), children who show aggressive behavior are proposed to have deficits in their cognitive SIP. Although ToM is not explicitly mentioned in the SIP model, the processes that are described for Step 2, *interpretation*, in the SIP model include ToM by nature, namely, making inferences about the perspectives of others and their intentions in the situation. Deficits in ToM have already been related to deficits in SIP empirically; for instance, deficits in ToM in preschoolers were predictive of deficits in SIP by hostile attributions of intent in school-aged children (Choe, Lane, Grabell, & Olson, 2013). Further, deficits in ToM are related to deficits in attributional processes and in forming a mental representation of a social situation (Teufel, Fletcher, & Davis, 2010); both are part of encoding and interpretation, Steps 1 and 2, of the SIP model. Further, it could be argued that ToM is also involved in other Steps of the SIP model, such as in the *response access or construction* (Step 4) or the *response decision* (Step 5), in which outcome evaluations play an important role. For both steps, the ability to infer the mental states of others and to use this knowledge to predict their behavior is very useful and could influence the performance of children in these steps of SIP. Empirically, it has been shown in a peer-coordination game that 6-year-olds adjusted their response decision by applying their ToM skills (Grueneisen, Wyman, & Tomasello, 2015). One reason why ToM as a construct was not included in the SIP model literally could be that the research on ToM was still at an early stage in 1996, when the SIP model was developed.

Another account states that children with conduct problems display a delayed or deviant understanding of intentions (Kenneth A Dodge & Frame, 1982) and emotions (Arsenio & Fleiss, 1996; Hughes et al., 1998). Understanding intentions and emotions of others is linked to cognitive and affective ToM (Shamay-Tsoory & Aharon-Peretz, 2007). Further, empathy and ToM are highly associated with one another according to Shamay-Tsoory and colleagues (e.g., Shamay-Tsoory, Harari, Aharon-Peretz, & Levkovitz, 2010), and it is assumed that deficits in both are also related (e.g., Shamay-Tsoory, Tomer, Berger, & Aharon-Peretz, 2003). So, ToM is a prerequisite to empathic responding (e.g., Sharp, 2008), which is proposed to be involved in an important inhibiting mechanism against violence and antisocial behavior (Blair et al., 2014). Empathy is critical for socialization and, as outlined above, deficits in empathy are proposed to be one etiological factor in conduct disorder. For instance, deficits in empathy are manifested in children who show higher levels of conduct problems in an impaired ability to recognize distress in peers or in lower levels of altruism toward the victim (e.g., Marsh et al., 2008). Further, the ability to take another's perspective, which is an aspect of ToM, is important for intentional moral conduct (e.g., Hoffman, 1987). Therefore, it is assumed that deficits in the ability to understand another's cognitive and affective perspectives are implicated in conduct problems (e.g., Anastassiou-Hadjicharalambous & Warden, 2008; Hare, 1970). One combining feature of these accounts is that they both involve activation of the amygdala. ToM performance is correlated with amygdala activation (e.g., Baron-Cohen et al., 1999), and deviant functioning of the amygdala has been found in youth with conduct disorder (e.g., Blair, 2005; Blair et al., 2014; Sebastian et al., 2012).

To sum up the two theoretical positions in light of the SIP model, deficits in children's social-cognitive processing are associated with aggressive behavior (e.g., Huesmann, 1998), and the amygdala theory of violent behavior proposes that deficits in ToM are related to deficits in empathy, which hinder an important inhibiting mechanism against violent and antisocial behavior (e.g., Blair et al., 2014; Sharp, 2008). However, the evidence on deficits in ToM in children with conduct problems is mixed. There is a large body of research that has confirmed a deficit in ToM in children with higher levels of aggressive behavior or conduct problems in early childhood (in 2-year-olds: Hughes & Ensor, 2006; in 5-year-olds: de la Osa, Granero, Domenech, Shamay-Tsoory, & Ezpeleta, 2016) and in middle childhood (in 7- to 10-year-olds: Anastassiou-Hadjicharalambous & Warden, 2008; in a clinical sample of 5- to 9-year-olds: Fahie & Symons, 2003; in 9- to 13-year-olds: Sharp, 2008); however, two studies found results that are difficult to interpret. First, hard-to-manage preschoolers performed worse in ToM tasks than controls (Hughes, Dunn, & White, 1998), but their performance in cognitive and affective ToM did not correlate with their conduct problems (Hughes, White, Sharpen, & Dunn, 2000). Second, Happé and Frith (1996) also obtained a contrasting result, because the 6- to 12- year-olds in their study mastered two simple ToM tasks but showed deficits on items that assessed ToM abilities relevant for everyday life. As a consequence, it was proposed that children with conduct problems may rather have an intact but deviant ToM – they call it a 'theory of nasty minds' (Happé & Frith, 1996, p. 395). So, children with conduct problems may use their intact ToM to carry out antisocial behavior like lying, cheating, or teasing others.

Along the same lines, another theoretical position proposes that aggressive children – instead of being 'socially inadequate – actually are more socially competent than their victims (Sutton, Smith, & Swettenham, 1999a). However, one has to note that this reasoning comes primarily from research on bullying. The definition of bullying relates only to a specific, repetitive form of aggressive behavior that includes a power asymmetry between perpetrator and victim (e.g., Olweus, 1999). This proposition has been confirmed in particular for ringleader bullies, who manipulate others to inflict suffering while they themselves avoid detection. In a study on 7- to 10-year-olds, the ringleader bullies scored higher than 'follower' bullies, victims, and defenders of the victim on a set of stories that assessed understanding of cognitions and emotions (Sutton, Smith, & Swettenham, 1999). One other study on 8- to 11-year-olds found support for Sutton et al.'s proposition, which is that victims showed more difficulties on a social cognition task, whereas bullies did not (Gini, 2006). Furthermore, Sutton et al. (1999) proposed that ToM skills may be particularly useful for relational or indirect ways of bullying. By applying their ToM ability, bullies can anticipate indirect ways that are most successful in harming their victims. Support for this proposition also comes from a longitudinal study, in which ToM at age 5 was positively associated with indirect aggression one year later, but only in children with low or average levels of prosocial behavior, while physical aggression was not related to ToM (Renouf et al., 2010). However, in this study, the term 'indirect' instead of 'relational' aggression was used, and the items refer to more indirect ways of harming others, which might be similar to indirect bullying. By contrast, another study on the relation between ToM and relational aggression found a negative link in preadolescents (Kokkinos, Voulgaridou, Mandrali, & Parousidou, 2016).

To sum up, a large body of studies has found evidence on ToM deficits in children with conduct problems, particularly research on physically aggressive behavior (e.g., Capage & Watson, 2001), but two studies found contrasting results (Happé & Frith, 1996; Hughes, White, Sharpen, & Dunn, 2000). Additionally, evidence on the relation between deficits in ToM and relational forms of aggression is less clear (e.g., Renouf et al., 2010); particularly research from the 'bullying tradition' found a positive relation (Sutton et al., 1999b). In conclusion, it is not yet clear whether ToM is positively or negatively associated with different forms of aggressive behavior, in particular relationally aggressive behavior, and conduct problems in children. In terms of SIP, deficits in social-cognitive processing are proposed in children who show aggressive behavior (Crick & Dodge, 1994), and related deficits in ToM are assumed to hinder a mechanism inhibiting violent and antisocial behavior in children (e.g., Sharp, 2008). However, other accounts propose that children might use their superior ToM abilities to manipulate others in terms of ringleader bullying (Sutton et al., 1999a), and that children who behave in antisocial ways might have ‘a theory of nasty minds’ (Happé & Frith, 1996, p. 395). Consequently, it is not clear whether deficits in ToM are a risk factor for conduct problems, including different forms of aggression.

Further, the number of studies that control for verbal ability is very limited. This seems important, because ToM and verbal abilities are strongly correlated in childhood (Milligan, Astington, & Dack, 2007). An additional unresolved question is the moderation of the relation between deficits in ToM and conduct problems by gender and age. As outlined above, females have been shown to display higher scores in ToM than males (e.g., Devine & Hughes, 2013), and the development of ToM continues throughout middle childhood and adolescence (e.g., Apperly et al., 2011; Vetter, Leipold, Kliegel, Phillips, & Altgassen, 2013). Further, gender differences in prevalence rates of conduct problems have also been reported, with boys displaying higher conduct scores than girls (e.g., Ravens-Sieberer et al., 2008). In addition, in terms of the developmental taxonomy of antisocial behavior, conduct problems are proposed to increase with the beginning of adolescence, where the gender difference in prevalence rates of conduct problems is narrowest (e.g., Moffitt et al., 2001). This implicates that the relation between deficits in ToM and conduct problems could be different in females and males, and in different age groups. However, previous evidence is very limited on potential moderation by both gender and age. One cross-sectional study provides preliminary evidence in which ToM was positively related to relational aggression only in younger children aged 7.6 years or younger, whereas the link was negative for children aged 9.6 years or older, and not significant in between (Gomez-Garibello & Talwar, 2015)

Another point that has been neglected in the research on deficits in ToM and conduct problems in children is the direction of influence. In light of the previously mentioned transactional models of the developmental trajectory of conduct problems (e.g., Frick & Morris, 2010; Beauchaine et al., 2015) and in accordance with the SIP model (Crick & Dodge, 1994), it can be assumed that reciprocal processes take place between ToM and conduct problems. This can be explained in terms of the SIP theory through the continuous retrieval of one's past social experiences, including aggressive behavior, stored as mental representations from the database during SIP in social interactions. In addition, it is known that a variety of social experiences can influence the development of ToM (for a review, see Hughes & Leekam, 2004); for instance, aggressive behavior at 2.5 years negatively predicted ToM skills four months later (Song, Volling, Lane, & Wellman, 2016). Further, also in the ontogenic processes model of externalizing disorders, it is assumed that a certain cognitive deficit, such as deficits in ToM, may elicit certain responses from others (i.e., the peers), which then feed back into the child's development and behavior (Beauchaine et al., 2015; Kimonis & Frick, 2010). However, most of the research until now has been cross-sectional, which does not allow for the testing of reciprocal effects, or longitudinal research that does not control for earlier levels of behavior (Renouf et al., 2010), which also prevents the detection of reciprocal processes. One other study investigated reciprocal processes between SIP and aggressive behavior in a developmental cascade (Lansford et al., 2010), and confirmed the propositions of the SIP model on reciprocal processes.

In sum, more research on the exact nature of the relation between ToM and conduct problems, including different forms of aggressive behavior, and longitudinal evidence on reciprocal processes between the two is needed. Future research should concentrate on the developmental periods of middle childhood and the transition to adolescence due to conflicting results in this age range as well as developmental considerations. Additional points include the control of ToM for its association with verbal abilities, and the moderating roles of gender and age.

2.3. Executive function

2.3.1. Unity and diversity of EF

There is great disagreement in the literature about the exact definition of EF, but most would agree that it is an umbrella term usually equated with purposeful, higher-order processes associated with the prefrontal cortex (Hughes, Graham, & Grayson, 2005). These cognitive processes are directly responsible for the conscious control of thought and action (Zelazo & Müller, 2012), and are a core component of self-control or self-regulation ability (Miyake & Friedman, 2012), which is vital for autonomous and adaptive psychological functioning (Séguin, Arseneault, & Tremblay, 2007). One reason for the different definitions of EF stems from the variety of research traditions examining mainly overlapping psychological processes from different angles. The different names each emphasize different aspects, for instance, cognitive control (Zelazo & Anderson, 2013), self-control (Moffitt et al., 2011) and effortful control (Rothbart & Rueda, 2005), to name a few of the terms that are essential for the developmental research on aspects of EF. New conceptualizations of EF also consider the motivational and emotional significance of a situation, which are proposed to influence the EF processes and the neural systems that support EF (Zelazo, Qu, & Kesek, 2010). Hot EFs are defined as top-down processes that operate in motivationally and emotionally significant situations in contrast to cool EF processes, which are supposed to operate in abstract, decontextualized situations that lack a significant affective or motivational component (e.g., Zelazo & Carlson, 2012). However, this differentiation will only be necessary in the additional analyses (see section 6), and apart from that the term “EF” is used only for the traditional conceptualization of EF that does not differentiate between cool and hot EF components. One thing all the different definitions have in common is that EF is related to intelligence, but cannot be equated with what we normally call intelligence (Zelazo & Anderson, 2013).

EF is not only difficult to define, but also difficult to measure (Jurado & Rosselli, 2007). One main reason for this is the so-called 'task-impurity problem.' This term refers to the problem that every measurement with a specific task proposed to measure a certain EF also shows a systematic degree of measurement error, which is not associated with the specific EF (Miyake et al., 2000). Miyake and colleagues (2000) proposed a – now widely accepted – solution by defining EF in a unity-and-diversity framework. They used structural-equation modeling to analyze a variety of tasks established to measure EF, which resulted in a structure of three latent factors that were identified as three core processes of EF, in short: updating, inhibition, and set-shifting. Thus, individual differences in EF were described as differences in updating (constant monitoring and manipulation of contents in working memory), inhibition (deliberate overriding of dominant or prepotent responses), and shifting or set-shifting (switching flexibly between tasks or mental sets), which is sometimes also called 'cognitive flexibility' (Miyake et al., 2000; Miyake & Friedman, 2012). These three EFs are moderately correlated with one another, which means they seem to tap a common underlying executive ability (unity aspect) but are also clearly separable (diversity aspect) (Miyake et al., 2000). This structure has been replicated in many studies and different age groups up to now, and also in the – for this thesis – relevant age groups of children aged 8 to 13 years (Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003) and in children 11 years old (Rose, Feldman, & Jankowski, 2011). In another study, the model was compared across age groups of 7-, 11-, 15-, and 21-year-olds, and mostly supported (Huizinga, Dolan, & van der Molen, 2006). Therefore, the framework by Miyake et al. (2000) was adapted for the present thesis to measure EF.

2.3.2. Development and measurement of EF

Historically, research on EF first focused on younger children, perhaps because the first break through and the most rapid improvements in the development of EF occur in preschool-aged children (e.g., Carlson & Moses, 2001; Garon, Bryson, & Smith, 2008). However, similarly to ToM, recent research has shown that the development of EF is ongoing throughout middle childhood and adolescence (e.g., Blakemore & Choudhury, 2006; Zelazo & Carlson, 2012). Thereby, the individual components of EF follow differential courses throughout development, including regressive – such as caused by synaptic pruning – and progressive – caused by myelination or synaptogenesis – stages of development (e.g., Best, Miller, & Jones, 2009). At the same time, the interindividual differences between persons seem to be relatively stable developmentally, which is also expressed by the high heritability currently attributed to EF (e.g., Miyake & Friedman, 2012).

There are a few problematic issues when describing the developmental trajectory of EF from childhood to adolescence, one of which is that usually complex tasks are used to measure EF in older children and adolescents to avoid ceiling effects (Best & Miller, 2010). However, applying complex tasks brings with it the already mentioned problem of 'task-impurity,' such as the tasks used to measure multiple EFs, and thus different researchers classify the same task as measuring different EFs. For instance, the dimensional change card sorting task is described as an inhibition task by some, and as a set-shifting task by others (Garon et al., 2008). Another topic that is directly related to the first one is that different tasks must be used across different age groups to avoid either ceiling or floor effects. This makes it difficult to directly compare the performance of different age groups with each other. One solution to these issues is to differentiate between simple and complex tasks when describing the development of a specific EF from childhood through adolescence (Garon et al., 2008; Best & Miller, 2010).

For working memory updating, simpler tasks require only the maintenance of information. An example is the forward digit span task, in which children have to repeat an increasing number of digits to which they have just listened before (Best & Miller, 2010). By contrast, more complex tasks require a greater degree of processing, such as the manipulation and maintenance of information, which requires coordination by a central executive like in the classic working memory model by Baddeley and Hitch (1974). A linear increase from ages 4 to 14 is reported for simple as well as complex tasks (Gathercole, Pickering, Ambridge, & Wearing, 2004). Additionally, the developmental course depends on the complexity of the task (Luciana, Conklin, Hooper, & Yarger, 2005), with the age of mastery depending on the degree of processing necessary for a particular task (Conklin, Luciana, Hooper, & Yarger, 2007). Thus, working memory updating seems to develop gradually with continued refinement in performance on complex tasks in adolescence, suggesting altogether a linear developmental trajectory from preschool through adolescence (Best & Miller, 2010).

In inhibition, complex tasks are distinguished from simple tasks based on whether they also afford working memory (Garon et al., 2008). Stroop tasks are an example of complex inhibition tasks, in which children must inhibit a prepotent or automatic verbal response and activate an alternative response. To do so, a rule has to be kept in working memory, which has to be either activated or inhibited to produce an alternative response. By contrast, simple tasks require only minimal amounts of working memory, such as delay of gratification tasks, in which a child chooses between a smaller immediate treat or a larger delayed reward. Inhibition shows rapid improvement in early childhood on a variety of response inhibition tasks, and by age 4 children are able to perform successfully on both simple and complex tasks (Best & Miller, 2010). Findings on further development after age 5 are rather mixed, with some support for additional significant improvements particularly from age 5 to 8 (e.g., Romine & Reynolds, 2005), and in complex tasks (e.g., Carlson, 2005). In sum, rapid early improvements are followed by slower improvements through adolescence, which probably involve quantitative refinements in accuracy instead of conceptual changes in inhibition (Best & Miller, 2010).

Set-shifting builds on the two other facets of EF, because shifting between two or more mental sets requires inhibition of previously activated mental sets and continuous updating of different rule sets (Best & Miller, 2010). Shifting usually means switching between two or more mental sets, each of which can contain several task rules, in contrast to inhibition, which only includes inhibition of a single, prepotent response (Crone, Somsen, Zanolie, & Van der Molen, 2006). The difference between simple and complex tasks is less clear in shifting, however more complex shifting tasks usually involve other cognitive processes, like metacognitive strategies (Best & Miller, 2010). Preschoolers ages 3 to 4 can successfully master simple shifting tasks, when rules are placed in a story context (e.g., Hughes, 1998), such as in the dimensional change card sort paradigm, where children have to sort cards that vary in two different dimensions. In the first phase, children have to sort cards based on one dimension (e.g., shape of figure), and are then required to shift and sort the cards based on another dimension (e.g., color of figure) (Frye, Zelazo, & Palfai, 1995; Zelazo, 2006). Between ages 5 and 6 further improvements occur in more complex shifting tasks (Luciana & Nelson, 1998), but are not interpreted as improvement in shifting per se. Instead, they can be seen as changes in other abilities, such as the ability to generalize learned rules to new settings (Best & Miller, 2010). In sum, preschoolers master simple shifting tasks, and learn to handle increasingly complex shifting tasks with time. Further development follows a protracted trajectory throughout adolescence in more complex shifting tasks that involve metacognitive abilities, such as monitoring one's errors; by middle adolescence, performance on these tasks typically reaches adult levels (Best & Miller, 2010).

Altogether, development in working memory updating, inhibition, and set-shifting follows a protracted course throughout middle childhood and adolescence, which emphasizes the investigation of age differences in research including EF in these age periods.

2.3.3. Gender differences in EF

Diamond (2013) puts the study of gender differences in EF on the list of future issues in her famous review on EF; indeed, surprisingly limited evidence on gender differences in EF has been found, with the sole exception of research on gambling and risk-taking, where a large number of studies investigate gender differences (for a review, see: van den Bos, Homberg, & de Visser, 2013). In general, one could expect gender differences in EF in middle childhood and adolescence due to the different development of the prefrontal cortex depending on sex (Blakemore & Choudhury, 2006). Among the few studies that investigated gender differences in inhibition, updating, and shifting, some found no difference between females and males (Huang-Pollock, Shapiro, Galloway-Long, & Weigard, 2017; in preschoolers: Rhoades, Greenberg, & Domitrovich, 2009; Welsh, Pennington, & Groisser, 1991); however, others did detect gender differences, such as males' superiority in speed (Brocki & Bohlin, 2004) and accuracy (e.g., in young children: Overman, 2004), with girls performing more slowly and making more errors. This result was in accordance with Becker, Isaac, and Hynd's (1987) study, which also found White Caucasian boys to react faster compared to White Caucasian girls, as well as to Black boys and girls. Brocki and Bohlin (2004) argue that this finding could be explained by a general gender difference in response styles, with girls tending towards a more cautious approach in the selection of an answer, which would lead to less accurate response patterns in girls under time pressure. Girls' tendency to be more conscientious was already found by Hagekull and Bohlin (1998) and by Gallagher (1998), who proposed that many gender differences in cognitive abilities are caused by cautiousness or another response bias. However, probably no single answer is sufficient to explain all gender differences in cognitive abilities.

By contrast, other studies found girls to perform better, for instance, in measures of verbal fluency (Reader, Harris, Schuerholz, & Denkla, 1994), inhibitory control (Berlin & Bohlin, 2002; Carlson & Moses, 2001), effortful control (Olson et al., 2005), self-control (Moffitt et al., 2011), and showing more growth in self-regulation in preschool (Montroy, Bowles, & Skibbe, 2016). Girls' superiority in measures including verbal or word fluency could be explained in light of females generally having better verbal abilities than males (e.g., Halpern & Lamay, 2000). Explanations for why females perform better in inhibitory control and delay of gratification tasks are more difficult to find. Some argue that girls have greater knowledge about strategies that are relevant for inhibition tasks; however, Silverman (2003) finds this explanation improbable. Altogether, very little is known about gender differences in working memory updating, set-shifting, and inhibition, and all cited authors agree that more study is needed on gender differences in EF.

2.3.4. Deficits in EF as a risk factor for conduct problems

Evidence and theoretical propositions on the relation between deficits in EF and conduct problems come from different lines of research paralleling the different research traditions on EF. In the developmental taxonomy of antisocial behavior, Moffitt (e.g., 1993) already outlined that neuropsychological deficits, particularly in language abilities and EF, are associated with childhood-onset antisocial behavior. Other research found that individuals who show adolescence-limited antisocial behavior also exhibit neuropsychological deficits, but quantitatively not as many as those in the life-course persistent group, who have shown childhood-onset antisocial behavior (Fairchild et al., 2013). This line of research was followed up by studies on self-control, which is defined as an umbrella construct integrating different concepts like executive function (EF) and self-regulation (e.g., Moffitt et al., 2011). Self-control was confirmed as an important predictor of aggression (Ayduk, Rodriguez, Mischel, Shoda, & Wright, 2007), and antisocial personality disorder and criminal conviction in young adults (Caspi, 1996). In the Dunedin longitudinal study, it could be confirmed that a gradient of self-control, from the lowest to the highest quintiles, measured in childhood predicted higher likelihoods of being convicted of a criminal offense by the age of 32, after accounting for social class origin and IQ (Moffitt et al., 2011). These findings suggest a relation between deficits in self-control and antisocial behavior across the whole spectrum of self-control.

Another account is the frontal-lobe hypothesis of emotional and behavioral regulation that states that deficits in EF are related to higher levels of physical aggression (Séguin, 2009). A famous historical example of how lesions in the frontal lobe can affect one's behavior and life is Phineas Gage. He worked for a railway company and suffered severe injuries to his frontal lobe in 1848. Phineas Gage survived the accident but was henceforth described as a totally different person. He had difficulty making plans, and was described as irritable, disrespectful, and profane (e.g., Goldstein & Naglieri, 2014).[1] The frontal-lobe hypothesis states that not only in individuals with confirmed lesions, like Phineas Gage, but also in people who show physical aggression and hyperactivity, cognitive-neuropsychological function is systematically poor (Séguin, 2009). Similarly, aggressive behavior in children has been interpreted in terms of impaired behavioral control (Hughes et al., 1998). Further, it is proposed that particularly individuals who show physical aggression from an early age possess impaired problem-solving strategies, which are seen as an important function of EF (Séguin & Zelazo, 2005). In this framework, it is proposed that EF is involved in social problem-solving, namely problem representation, planning, execution, and evaluation (Zelazo, Carter, Reznick, & Frye, 1997). Deficits are hypothesized to possibly occur in all stages of the problem-solving process, for example in representing a problem adequately, manifested in the failure to plan a solution, the failure to use rules, and difficulties with evaluating actions, as well as the failure to learn from negative consequences (Séguin & Zelazo, 2005).

Difficulties and deficits in SIP are also the focus of the model by Crick and Dodge (1994) to explain aggressive behavior in children, as outlined above (see 2.1.6). Because of the complexity of social situations, it is cognitively demanding to accurately process all relevant information (Zadeh, Im-Bolter, & Cohen, 2007). This requires higher-order cognitive

[1] New analyses of the injuries and behavior of Phineas Gage reveal that most of the older interpretations of the case are probably wrong; see Kean's book (2014) for a confrontation of the different scientific positions.

capacities, such as EF, that are assumed to be less adequate in children who show higher levels of aggressive behavior and conduct problems (Huesmann, Eron, & Yarmel, 1987; de Castro & van Dijk, 2018). Therefore, the deficits in SIP that are typical for aggressive children might be influenced by deficits in children's EF. For instance, children with externalizing behavior problems exhibit lower levels of working memory and also show deficits in their ability to encode and represent the problem in a given situation (Steps 1 and 2), select a goal for a specific social situation (Step 3), and evaluate and choose a response (Step 5) (Zadeh et al., 2007). Further, children with impaired inhibition evaluate aggressive responses more positively, and select aggressive responses more often (Van Nieuwenhuijzen et al., 2017). In addition, children's inhibition and planning abilities interact with their hostile encoding and attribution biases in the prediction of reactive and proactive aggression (Ellis, Weiss, & Lochman, 2009).

Research on effortful control contributes to evidence on the relation between deficits in EF and aggressive behavior. Effortful control is assumed to contribute to children's externalizing behavior by interfering with the processing of relevant information and the modulation of emotional experience and behavior (Eisenberg, Spinrad, & Eggum, 2010). Effortful control is defined as the ability to inhibit a dominant, prepotent response to perform a subdominant, less salient response and to correct failures, and continues to develop from an early age (Rothbart & Rueda, 2005). The conceptual differences and similarities between EF and effortful control strategies are still under discussion, but they seem to be both strongly associated and overlapping constructs (Bridgett, Oddi, Laake, Murdock, & Bachmann, 2013).

Empirical support for the relation between deficits in EF or effortful control and conduct problems, including aggressive behavior, comes from a large body of research in early childhood (e.g., Spinrad et al., 2007). Examples include meta-analyses among preschoolers (Schoemaker, Mulder, Deković, & Matthys, 2013) and in middle childhood (e.g., Eisenberg et al., 2004; Huang-Pollock, Shapiro, Galloway-Long, & Weigard, 2017). An older meta-analytic review (Morgan & Lilienfeld, 2000) and a more recent meta-analysis (Ogilvie, Stewart, Chan, & Shum, 2011) confirmed the association between deficits in EF and conduct and/or oppositional defiant disorder across the lifespan, although, in adolescence, only one recent study examined and found a link between error scores of EF tasks and conduct problems in a clinical sample (Dolan & Lennox, 2013). Other evidence on the relation in adolescence comes mainly from older studies on adolescents who were incarcerated or listed in criminal records (see Morgan & Lilienfeld, 2000). Considering longitudinal relations, evidence comes particularly from research on effortful control (e.g., Eisenberg et al., 2004); for instance, children with deficits in effortful control at age 3 showed an increased risk for a chronic pattern of higher levels of externalizing behavior problems through to age 11 (Olson, Choe, & Sameroff, 2017).

However, evidence on the association between deficits in EF and specific forms of aggression, particularly relational forms, is more limited. Most studies until now have focused on physical aggression (e.g., Schoemaker et al., 2013), which is in line with the frontal-lobe hypothesis (Séguin, 2009) and the problem-solving account (Séguin & Zelazo, 2005), both of which propose that deficits in EF are related to physical aggression in particular. However, as relational forms of aggression become more prominent after early childhood (e.g., Côté et al., 2007), they should also be considered particularly in studies situated in middle childhood and adolescence. Studies that included both forms of aggression revealed mixed findings. Some found evidence that deficits in EF were related to physical, but not relational aggression (O'Toole, Monks, & Tsermentseli, 2017), whereas other research found negative associations to both physical and relational aggression (McQuade, Murray-Close, Shoulberg, & Hoza, 2013; Poland, Monks, & Tsermentseli, 2016). Moreover, one study only found a negative association to relational, but not to physical aggression (Granvald & Marciszko, 2016), and another found no associations to either physical or relational aggression after controlling for symptoms of attention deficits/hyperactivity disorder (Diamantopoulou, Rydell, Thorell, &

Bohlin, 2007). Regarding the functions of aggression, little is known, because few studies have examined the link between reactive and proactive aggression and EF, particularly in middle childhood. Studies that examined the link found that deficits in EF were associated with reactive, but not with proactive aggression (Ellis, Weiss, & Lochman, 2009; Rathert, Fite, & Gaertner, 2011; White, Jarrett, & Ollendick, 2013). This can be explained by the mediating role of anger, which is the focus of the next section (2.3.5).

In sum, different lines of research have found evidence for a relation between deficits in EF and different kinds of externalizing behavior, particularly physically aggressive behavior and conduct problems. However, research on the association between deficits in EF and specific forms and functions of aggression apart from physical aggression is limited and, particularly considering relational aggression, mixed (e.g., Diamantopoulou et al., 2007; Granvald & Marciszko, 2016; McQuade et al., 2013). Furthermore, most of the research on the relation between deficits in EF and conduct problems has focused on preschool-aged or school-aged children, and less is known about this relation in early and middle adolescence, as well as whether there are differences based on participants' age. In addition, as the meta-analysis by Ogilvie et al. (2011) shows, only a minority of studies have included females as participants, and consequently, very little is known about the relation between deficits in EF and conduct problems in females.

2.3.5. Anger reactivity as a mediator in the relation between EF and aggressive behavior

As explained above (see 2.3.3), anger and aggressive behavior in children are closely related, as anger seems to mediate between social-cognitive processes and aggressive behavior (e.g., de Castro et al., 2005). Further, EF is involved in the regulation of negative affect, for instance in toddlers (Putnam, Rothbart, & Gartstein, 2008) and preschool children (Carlson & Wang, 2007). Research on irritability – an increased proneness to anger – found that higher cognitive control, a concept closely related to EF, seems to buffer against maladaptive outcomes of chronic anger, like adult antisocial personality (Hawes et al., 2016). In addition, deficits in EF are associated with higher levels of negative affect, including anger (Bridgett et al., 2013; Gagne & Hill Goldsmith, 2011; Healey, Marks, & Halperin, 2011). Further, children with higher rates of irritability showed deficits in processes related to EF, namely, aberrant processing of emotional stimuli and impaired context-sensitive regulation (Leibenluft & Stoddard, 2013), as well as neural dysfunction in error monitoring, reward processing, and emotion regulation (Perlman et al., 2015). Children who are prone to anger were also found to show more aggressive behavior (e.g., Arsenio, Cooperman, & Lover, 2000; Eisenberg et al., 1999; Olson et al., 2005; Wakschlag et al., 2015), in particular reactively, but not proactively aggressive behavior (for a review, see Hubbard, Romano, McAuliffe, & Morrow, 2010). Consequently, deficits in EF may increase the experience of anger, and as such, anger may mediate between social-cognitive processes, that is, EF, on the one side, and aggressive behavior, particularly reactive aggression, on the other.

2.4. The relation between EF and ToM

ToM and EF are often studied in tandem because of their relative interdependence. A recent meta-analysis reported a correlation of $r = .38$ in 3- to 6-year-old children from 15 countries (Devine & Hughes, 2014). At least four different accounts that explain this correlation developmentally are discussed in the literature (e.g., Moses & Tahiroglu, 2010). First, a common underlying cognitive component, like *hierarchical rule use*, could be responsible for the stable relation between ToM and EF (e.g., Frye, Zelazo, & Burack, 1998). A second possibility is that some component of ToM is required for the development of EF. Perner and colleagues (2002) argue that *meta-representational capacities* central to ToM are required for certain executive abilities, like inhibitory control (Perner, Lang, & Kloo, 2002). Opposed to this view is a third opinion that EF might be a requirement for children to *express* their ToM abilities. Children might need some inhibitory ability to distance themselves from their own point of view or cognitive flexibility to shift their perspective away from their own point of view to another's, which is why they are also called *executive performance* and *competence* accounts (e.g., Apperly, Warren, Andrews, Grant, & Todd, 2011). Finally, a fourth account determines that EF skills are necessary for the *emergence* of ToM, that is, the acquisition of mental-state concepts themselves (e.g., Carlson, Claxton, & Moses, 2015). Further, all named factors could also contribute to at least some extent to the correlation between EF and ToM (Moses & Tahiroglu, 2010), or social processes could be responsible for both ToM and EF development (e.g., Hughes, 1998).

A meta-analysis confirms that in early childhood earlier EF predicts later ToM measured via false belief tasks, but not vice versa, which supports the *emergence* and *expression* accounts (Devine & Hughes, 2014). Research on the relation between EF and ToM in adolescence is very limited, but initial evidence supports a link between EF and affective ToM (Vetter, Altgassen, Phillips, Mahy, & Kliegel, 2013). A weak to moderate link between EF and ToM in middle childhood has been confirmed in a few cross-sectional studies (Bock, Gallaway, & Hund, 2014; Cantin, Gnaedinger, Gallaway, Hesson-McInnis, & Hund, 2016; Hansen Lagattuta, Sayfan, & Harvey, 2014; Huyder, Nilsen, & Bacso, 2017), across different measures of EF and ToM (e.g., Kouklari, Thompson, Monks, & Tsermentseli, 2017) and across different cultures (Wang, Devine, Wong, & Hughes, 2016). Considering longitudinal relations in middle childhood, evidence so far supports the idea that earlier EF, in particular updating and attention shifting, predicts later ToM, which is in line with the *executive performance* and *competence* accounts (Austin, Groppe, & Elsner, 2014; Lecce, Bianco, Devine, & Hughes, 2017). However, another longitudinal study in middle childhood in children over 4 years of age (ages 6 to 10) found no longitudinal relations between EF and ToM, but this could be a consequence of using different tasks at the two waves to ensure age-appropriate task demands and avoid ceiling or floor effects (Devine, White, Ensor, & Hughes, 2016).

Altogether, EF and ToM seem to be interrelated throughout early and middle childhood to adolescence, which implies that they have to be studied in tandem. However, more research is needed on the link, particularly longitudinal evidence, and age-appropriate task demands must be ensured.

2.5. Research questions for the thesis and introduction to the studies

Evidence from behavioral and fMRI studies confirms that ToM development is ongoing from early childhood through adolescence (see e.g., Miller, 2009), as is development of EF (see Best & Miller, 2010). The described ongoing cognitive development parallels simultaneous developments at the social and emotional levels, and is in line with new arising developmental tasks due to the transition to elementary school in middle childhood (e.g., Epps & Smith, 1984). Some examples for developmental challenges in middle childhood are quantitative and qualitative changes in social interactions, such as beginning and maintaining intimate friendships (e.g., Hartup, 1984; Masten et al., 1995) and developing strategies for controlling and managing one's behavior in the school setting (e.g., Markus & Nurius, 1984).

At the same time, the use of physically aggressive behaviors decreases and the use of relationally aggressive behaviors increases (e.g., Côté et al., 2007). Regarding conduct problems, the prevalence of conduct symptoms increases with adolescence (e.g., Moffitt et al., 2001), which is in line with the developmental taxonomy of conduct disorder by Moffitt (1993). In addition, the transition to adolescence brings with it new developmental tasks, initiated by the hormonal and biological changes related to the beginning of puberty (e.g., Steinberg & Morris, 2001). All these considerations emphasize why the developmental periods of middle childhood and early adolescence are particularly interesting for the study of the relations between ToM, EF, and conduct problems.

Theoretical Background

The SIP model by Crick and Dodge (1994) serves as a theoretical background for this thesis. Deficits in ToM can be integrated into this model, and are related to deficits in SIP (e.g., Choe et al., 2013). Further, both are implicated in children with aggressive behavior (e.g., Capage & Watson, 2001; Lansford et al., 2010). Other accounts propose that deficits in ToM are involved in hindering an important inhibiting mechanism that prevents violent and antisocial behavior, which is also seen as an etiological factor in conduct problems in children (e.g., Blair et al., 2014). However, as outlined above, evidence on these relations is mixed. Further, only little research until now has concentrated on the developmental periods of middle childhood and early adolescence. The first research aim is to shed light on these contrasting results, and to determine whether deficits in ToM are a risk factor for conduct problems in this age range (I; path a, b, Figure 2). This research aim is examined in the first and third studies included in this thesis, although it is investigated under different research aspects, as explained in the sections introducing Studies 1 and 3 below.

The second research aim is to determine whether deficits in EF are also a cognitive risk factor for conduct problems in middle childhood and early adolescence (II). Earlier theoretical accounts have focused on the relation between deficits in EF and physical aggression, such as the problem-solving account by Séguin and Zelazo (2005) and the frontal lobe hypothesis (e.g., Séguin, 2009). Because of the complexity of social situations, it is cognitively demanding to accurately process all relevant information (Zadeh, Im-Bolter, & Cohen, 2007). This requires higher-order cognitive capacities, like EF, that are assumed to be less adequate in children who show higher levels of aggressive behavior and conduct problems (Huesmann, Eron, & Yarmel, 1987; de Castro & van Dijk, 2018; Van Nieuwenhuijzen et al., 2017). As a consequence, the deficits in SIP that are typical for aggressive children might be influenced by deficits in children's EF. Further, most studies until now have focused on physical aggression. Less is known about the relation between deficits in EF and relational aggression, reactive and proactive aggression, and conduct problems in adolescent non-clinical samples. Consequently, the aim here is to clarify whether deficits in EF are a risk factor for conduct problems, including different forms and functions of aggression (path d), in a community sample in the developmental periods of middle childhood and early adolescence.

Furthermore, most research up to this point has examined deficits in ToM and EF as risk factors for conduct problems in separate studies despite their interrelatedness throughout development as well as evidence for a dependence of ToM performance or competence on EF (e.g., Apperly et al., 2011). Some research has been conducted with younger children, but evidence is mixed on the cross-sectional links between EF, ToM, and conduct problems. Some evidence hints at a direct effect of deficits in EF on conduct problems (e.g., O'Toole et al., 2017), but not all (e.g., Hughes & Ensor, 2006). And one study has suggested that deficits in EF and conduct problems are not only directly linked, but also indirectly through impaired ToM (Hughes et al., 1998). Therefore, the third study examines whether deficits in ToM and EF are risk factors for conduct problems when they are studied in tandem in this age range, which is the third research question (III; paths a, c, d).

All three studies were embedded in a large, longitudinal study on intrapersonal risk factors for psychopathology in children and adolescents ('Potsdamer Intrapersonale Entwicklungsrisiken Studie – PIER-Studie') conducted at the University of Potsdam. In this study, three time points were measured, all three of which were included in Studies 1 and 2, while only the third measurement point was included in Study 3. The distance between the first two time points was about 9 months, and the distance between the second and the third was about 2 years. The community sample consisted of 1,657 children ages 6 to 13, with an equal distribution of female and male children (52.1% girls). The data collection took place at local schools in Brandenburg, Germany, which included rural and urban areas. Additional data were collected from teachers and parents of the children by questionnaires.

Theoretical Background

Study 1 (Holl, Kirsch, Rohlf, Krahé, & Elsner, 2018)

One reason for the contrasting results on the relation of ToM and conduct problems could be the fact that different forms of aggression – physical and relational – were not considered as separate domains in the study of the association of ToM and conduct problems. Aggressive behavior is considered as one important component of conduct problems (e.g., Lahey et al. 2000), but physical and relational aggression has been shown to relate differently to ToM. Evidence on physical aggression and ToM mainly found a negative relation (e.g., Capage & Watson, 2001), however, relational forms of aggression could also be associated positively to ToM (e.g., Renouf, 2010). Consequently, in the first study, the relation of physically and relationally aggressive behavior and ToM were examined over the course of middle childhood through to the transition to early adolescence.

Thereby, we adopted a developmental perspective through our longitudinal design, which allowed for reciprocal processes, as it was already proposed in the SIP model (Crick and Dodge, 1994), and is requested by newer transactional models of externalizing psychopathology (e.g., Beauchaine et al., 2015). This means that we not only hypothesized effects from ToM to aggressive behavior (Figure 2, path a), but also from aggressive behavior to ToM (path b) over time. Additional research aims were the investigation of gender and age differences in the examined relations.

Study 2 (Rohlf, Holl, Kirsch, Krahé, & Elsner, 2018)

The second study focused on the relation between deficits in EF and forms and functions of aggressive behavior over the course of middle childhood, and the mediating role of anger in this relation. As previous research has confirmed, deficits in EF are related to externalizing behavior, which includes aggressive behavior, in preschool-aged children (e.g., Schoemaker et al., 2013). Different theoretical accounts refer to this association, for instance, the frontal-lobe hypothesis of aggression (e.g., Séguin, 2009), a problem-solving framework (Séguin &

Zelazo, 2005), and the developmental taxonomy of conduct disorder by Moffitt (1993). And is also proposed that deficits in SIP could be related to deficits in EF in children with conduct problems (de Castro & van Dijk, 2018; Zadeh, Im-Bolter, & Cohen, 2007). However, evidence on this relation in middle childhood and in the transition to adolescence is less clear, in particular when it comes to the forms and functions of aggression in non-clinical samples and in females. As a consequence, we examined the relation of deficits in EF and aggressive behavior in two models (Figure 2, path d). In one, we investigated the relation between deficits in EF and physical and relational aggression, and in the other between deficits in EF and reactive and proactive aggression. Further, we investigated the mediating role of anger between deficits in EF and reactively aggressive behavior, because anger has been shown to be related to aggression, particularly to reactive aggression (e.g., Ellis et al., 2009; Rathert, Fite, & Gaertner, 2011). Additionally, we examined the potential moderating roles of gender and age in our models.

Study 3 (Holl, Vetter, & Elsner, 2018, submitted)

In the third study, the relation of ToM and conduct problems was investigated at the transition from middle childhood to early adolescence. Research until now has primarily considered the relations of ToM, EF and conduct problems in separate studies and analyses. As ToM and EF are interrelated trough development, it deemed important to study these two cognitive domains in tandem in the relation to conduct problems, which was the research aim of the third study. Thereby, we investigated the impact of ToM (Figure2, path a), and of EF on conduct problems (path d), and the relation of ToM and EF (path c) in middle childhood and at the transition to adolescence. Thereby, we examined direct and indirect effects of EF on conduct problems, because as depicted in Figure 2, an indirect path from EF via ToM to conduct problems was supposed due to the theoretical propositions on the relation of EF and ToM. This model could suggest a mediational process; however, we found no independent, direct effect of EF on conduct problems, and therefore the requirement for a mediation was not fulfilled. In addition, in this study the analysis was purely cross-sectional, and to determine a mediational process throughout development, one needs at least three measurement points over time (Maxwell, Cole, & Mitchell, 2011). Considering previous evidence on gender differences in ToM, EF, and conduct problems, and in dependence of the large age-range in our sample, we examined, if the relations in our model were moderated by gender and age.

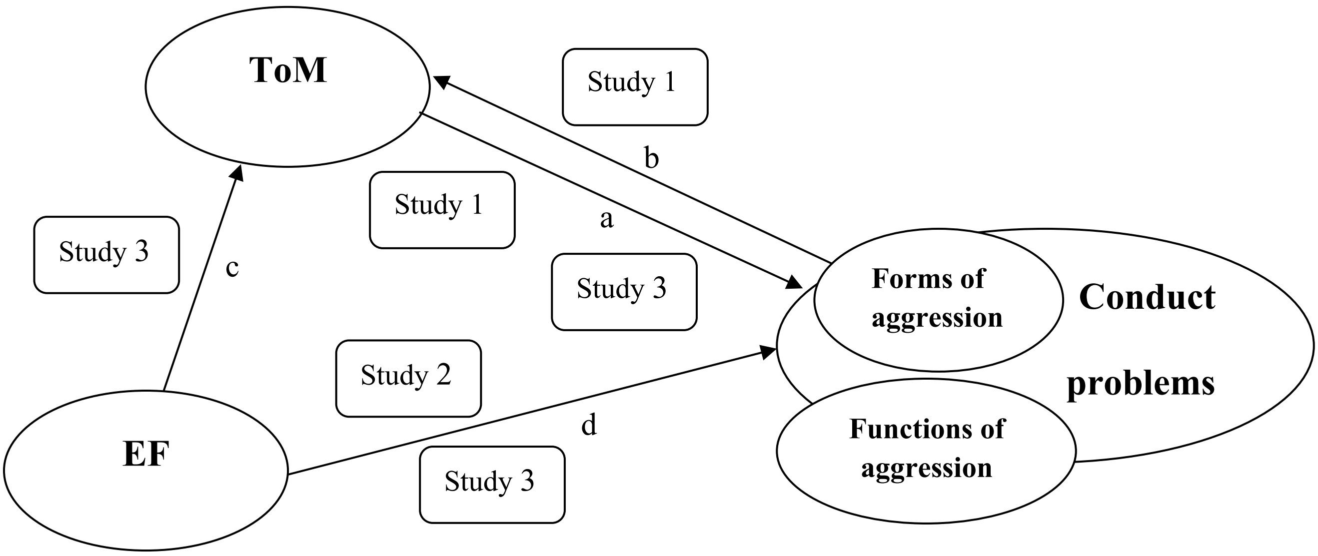


Figure 2. The examined relations in the thesis and studies

The research questions for the thesis are as followed:

I) Are deficits in ToM a risk factor for conduct problems, including different forms of aggressive behavior, from middle childhood throughout early adolescence?

II) Are deficits in EF a risk factor for conduct problems, including different forms and functions of aggressive behavior, from middle childhood throughout early adolescence?

III) Are deficits in EF and ToM risk factors for conduct problems from middle childhood throughout early adolescence, when they are studied in tandem?

Additional research aims were the moderating role of gender and age differences in the examined relations.

The research questions were examined in the three studies, which are displayed in the next three sections.

3. Study 1

Holl, A. K., Kirsch, F., Rohlf, H., Krahé, B., & Elsner, B. (2018). Longitudinal reciprocity between theory of mind and aggression in middle childhood. *International Journal of Behavioral Development, 42(2)*, pp. 257-266.

http://doi.org/10.1177/0165025417727875

Study 1

Abstract

Theory of mind is one of the most important cognitive factors in social information-processing, and deficits in theory of mind have been linked to aggressive behavior in childhood. The present longitudinal study investigated reciprocal links between theory of mind and two forms of aggression – physical and relational – in middle childhood with three data waves over 3 years. Theory of mind was assessed by participants' responses to cartoons, and physical and relational aggression were assessed through teacher reports in a community sample of 1657 children (mean age at Time 1: 8 years). Structural equation modeling analyses showed that theory of mind was a negative predictor of subsequent physical and relational aggression, both from Time 1 to Time 2 as well as from Time 2 to Time 3. Moreover, relational aggression was a negative predictor of theory of mind from Time 1 to Time 2. There were no significant gender or age differences in the tested pathways. The results suggest that reciprocal and negative longitudinal relations exist between children's theory of mind and aggressive behavior. Our study extends current knowledge about the development of such relations across middle childhood.

Study 1

Introduction

Aggressive behavior is one of the most pressing social problems in our world today, because of its detrimental effects on the lives and health of human beings (Dahlberg & Krug, 2002). Already in childhood, aggression has negative consequences, for instance, in the form of maladjustment of both children who show aggressive behavior (e.g., Crick et al., 2006) and children who are the targets of that behavior (e.g., Crick et al., 2001). Social cognition in general and theory of mind (ToM) in particular have been recognized as core processes involved in the development of aggressive behavior (Huesmann, 1998). However, whether ToM is negatively or positively linked to aggression and whether ToM is a cause or even a consequence of aggression are issues that have not yet been sufficiently addressed. Thus, the present study aimed at examining potential reciprocal relations between ToM and aggressive behavior in middle childhood.

Aggressive behavior is defined as "any form of behavior directed toward the goal of harming or injuring another live being who is motivated to avoid such treatment" (Baron & Richardson, 1994, p. 7). Harming others can be done by physical means, for example by hitting or kicking, known as *physical aggression* (PHY-A; e.g., Crick, Casas, & Ku, 1999), or by damaging their friendships or feelings of inclusion in the peer group, for example by spreading rumors, known as *relational aggression* (REL-A; e.g., Crick & Grotpeter, 1995). Some authors use the terms indirect or social aggression instead of REL-A. However, all three constructs essentially refer to the same form of aggression (Archer & Coyne, 2005), we refer to REL-A, even when the cited authors used an alternative term. Regarding the developmental course of the forms of aggression, the majority of children, girls in particular, reduce their use of PHY-A (Côté, Vaillancourt, Barker, Nagin, & Tremblay, 2007) and tend to increase their use of REL-A (Crick et al., 1999) during middle childhood.

There are typical gender differences in the two forms of aggression, with boys engaging in PHY-A more often than girls (e.g., Lansford et al., 2012). Theoretical explanations for this gender difference in PHY-A include hormonal, evolutionary, and social role accounts (see Krahé, 2013, for review). With regard to REL-A, several studies have found that girls engage in REL-A more often than boys (e.g., Ostrov, Murray-Close, Godleski, & Hart, 2013). One explanation is that girls tend to have more dyadic relationships than boys, which seems to fuel REL-A (Murray-Close, Ostrov, & Crick, 2007). However, other studies revealed either no gender differences in REL-A (e.g., Lansford et al., 2012) or higher scores for boys (e.g., Henington, Hughes, Cavell, & Thompson, 1998). These mixed findings highlight the importance of conducting further research on the developmental courses of PHY-A and REL-A during middle childhood, thereby including both forms of aggression and differentiating developmental patterns for girls and boys.

The Social Information-Processing (SIP) Theory of Aggression

Crick and Dodge's (1994) social information-processing (SIP) model proposes that deficits and biases in processing social information, especially the misinterpretation or neglect of important social cues, can cause social maladjustment in children, including aggressive behavior. SIP is conceptualized as a sequence of six processing stages: (1) encoding and (2) interpretation of social cues, (3) clarification of individual goals, (4) access to or construction of potential behavioral responses, (5) decision for a response, and (6) behavioral enactment. This stepwise approach has helped to identify differences in SIP patterns in children who show higher as compared to lower levels of aggression. For instance, aggressive children tend to attribute hostile intentions to their peers in ambiguous social situations more often than non-aggressive children (de Castro, Veerman, Koops, Bosch, & Monshouwer, 2002, SIP Step 1 and 2). Furthermore, aggressive children generate more aggressive responses than their non-aggressive peers (de Castro, Merk, Koops, Veerman, & Bosch, 2005, SIP Step 4).

Study 1

In terms of the chronification of aggressive behavior, Crick and Dodge (1994) assumed that the relation between social experiences and SIP is reciprocal. They proposed that mental representations of social experiences are stored in long-term memory, in a so-called database, and are then accessed in each of the postulated SIP steps. Thus, social knowledge is retrieved and, in turn, shapes behavior at each step. Consequently, social cognition influences the occurrence of aggressive behavior, and, in turn, memories of past aggressive behavior influence social cognition. In one of the few prospective studies that have examined this reciprocal link, Lansford, Malone, Dodge, Pettit, and Bates (2010) demonstrated a dynamic developmental cascade in which SIP and aggressive behavior influenced one another over the course of middle childhood. Crick and Dodge (1994) proposed to investigate the impact of social experiences, such as aggressive behavior, on social cognitions. This can foster the understanding of possible mechanisms by which children develop outcome expectations for specific social situations, which, in turn, children may use to reach a response decision (SIP Step 5).

ToM, which enables individuals to "predict what others are going to do on the basis of their desires and beliefs" (Frith & Frith, 2010, p. 165), is involved in at least some steps of the SIP model (Harvey, Fletcher, & French, 2001). ToM is typically distinguished into two facets: Making inferences regarding others' beliefs, intentions, or desires, refers to *cognitive* ToM, and inferring others' emotions refers to *affective* ToM (e.g., Shamay-Tsoory, Harari, Aharon-Peretz, & Levkovitz, 2010). In the last decade, behavioral (Devine & Hughes, 2013) as well as neuroscientific studies (Vetter, Weigelt, Döhnel, Smolka, & Kliegel, 2014) confirmed the further development of ToM after the accomplishment of understanding of false belief during preschool age. The development of cognitive ToM (Devine & Hughes, 2013) as well as the understanding of advanced affective ToM (Vetter, et al., 2014) are assumed to continue throughout middle childhood and adolescence.

Both facets of ToM are assumed to influence at least some of the steps postulated in the SIP model. For instance, ToM plays an important role in attributional processes and in forming a mental representation of a social situation (Teufel, Fletcher, & Davis, 2010), both of which are part of encoding and interpretation of social cues (SIP Steps 1 and 2). Furthermore, ToM seems to play a role in deciding for a response (SIP Step 5), as has been shown in a peer-coordination game, in which 6-year-old children adjusted their response decisions by applying their ToM skills (Grueneisen, Wyman, & Tomasello, 2015).

Theory of Mind and Aggressive Behavior

Numerous studies have examined the relation between ToM and aggressive behavior in children. However, most of these studies yielded mixed results, were situated in preschool age, were cross-sectional, or did not control for earlier levels of behavior. Some studies supported the 'social skills deficit view' described above, which assumes that deficits in ToM co-occur with a pattern of SIP that fosters aggressive behavior (Crick & Dodge, 1994). For instance, deficits in ToM were correlated with behavioral problems in 2-year-olds (Hughes & Ensor, 2006), as well as in a clinical sample of 5- to 9-year-olds (Fahie & Symons, 2003). Similarly, preschool-aged children with superior ToM were rated as being less aggressive toward their peers (Capage & Watson, 2001).

However, in contrast to the social skills deficit view, Sutton, Smith, and Swettenham (1999a) proposed that children who show aggressive behavior toward their peers, instead of being 'socially inadequate', are more socially competent than their victims. This reasoning comes primarily from research on bullying, as a specific form of aggressive behavior that is ongoing and involves a power asymmetry between perpetrator and victim (Olweus, 2013). In children with low ToM abilities, bullying may result from deficits in SIP (e.g., misinterpretation of social cues), as described above. However, bullies also may have superior ToM abilities that enable them to manipulate others, which was demonstrated particularly for ringleader-bullies in middle childhood (Sutton, Smith, & Swettenham, 1999b). Sutton et al. (1999a) further argue that ToM abilities may be particularly relevant for the use of relational forms of bullying, because children with higher ToM abilities are better able to anticipate the victim's reactions to REL-A, and by which means they can best harm their victim. Although this assumption was formulated for relational bullying, it may be applicable to REL-A in general, because most relationally aggressive acts require the involvement of others. In line with this reasoning, in one of the few longitudinal studies addressing the link between ToM and aggression, 5-year-old children's ToM skills were positively related to REL-A one year later, but only in children with average or low levels of prosocial behavior (Renouf et al., 2010). In this study, PHY-A was unrelated to ToM.

Because most previous studies were cross-sectional (Gomez-Garibello & Talwar, 2015), or did not control for earlier levels of behavior, not much is known about the direction of effects between ToM and different forms of aggression, PHY-A and REL-A, particularly in middle childhood. There is some evidence that ToM influences later aggression, for instance in preschool age (Renouf et al., 2010). However, the reverse path is also possible, as has been found in preschoolers, whose aggressive behavior at 2.5 years negatively predicted their ToM skills 4 months later (Song, Volling, Lane, & Wellman, 2016). The SIP model proposes reciprocal relations between social cognition (i.e., ToM) and social behavior, based on a continuous retrieval of stored mental representations of social experiences from one's database (Crick & Dodge, 1994). In addition, not much is known about potential changes in this link during middle childhood, when the further development of ToM abilities (e.g., Hughes & Devine, 2015) coincides with changes in PHY-A and REL-A (e.g., Côté et al., 2007). After the transition to school, a variety of social experiences, such as involvement in aggressive behavior, could influence the further development of ToM (see Hughes & Leekam, 2004, for a review on the links between ToM and social relationships). Yet, to our knowledge there are no longitudinal studies that examined both the potential path from ToM to PHY-A and REL-A, and the reverse path from these aggression forms to ToM at the same time over the course of middle childhood.

Additionally, the role of gender and age as potential moderators of the links between ToM and aggression requires further study. As explained above, girls and boys differ in their expression of REL-A and PHY-A (e.g., Lansford et al., 2012), therefore the links to ToM could also differ between both genders. Until now, the very few studies that considered gender differences in the relation between ToM and aggression produced inconsistent results (e.g., Kokkinos, Voulgaridou, Mandrali, & Parousidou, 2016; Renouf et al., 2010). With regard to age, only few studies have considered age as a moderator in the relations between ToM and aggression. As this link was mainly examined in children in preschool age, even less is known about possible age differences in middle childhood. One study found a moderation by age in the relation between REL-A and ToM in South-American children (Gomez-Garibello & Talwar, 2015).

The Current Study

In our three-wave longitudinal study, we examined the reciprocal relations of ToM with PHY-A and REL-A in middle childhood in a large community sample over a period of 3 years. At the first data wave (T1), participants had a mean age of 8 years, the second wave (T2) was conducted about 9 months later, and the third wave (T3) another 24 months later. Using structural-equation modeling, we investigated the reciprocal relations while controlling for earlier levels of ToM and partialling out stable between-person differences in aggressive behavior, respectively.

On the basis of the SIP model (Crick & Dodge, 1994) and in line with earlier empirical evidence (e.g., Capage & Watson, 2001), we expected a negative path from ToM to subsequent PHY-A. Regarding the path from ToM to later REL-A, the SIP model also suggests a negative path, whereas other research, particularly in the bullying tradition, suggests a positive path (e.g., Renouf et al., 2010). We examined these competing predictions in our analyses. Additionally, in line with the SIP approach, we postulated prospective paths from aggressive behavior to ToM. Past social experiences are stored in long-term memory and should be retrieved in every step of SIP. Since ToM is involved in at least some of the SIP steps, it might be influenced by earlier aggression. We therefore hypothesized negative paths from both PHY-A and REL-A to subsequent ToM.

In addition, we explored possible gender and age differences in the reciprocal associations between aggression and ToM. This was deemed important, because the inconsistent findings across previous studies point toward the need for further investigations to clarify the potential moderating influence of gender and age on the paths between ToM and aggression. To examine whether the proposed relations could be confirmed for all subsamples of our study, we conducted separate multi-group analyses with the gender groups and with three age groups, respectively.

Method

Participants

At the first data wave (T1) 1,657 children (52.1% girls) between 6 and 11 years (M = 8.36, SD = 0.95, range 6.23-11.33) participated. At the second wave (T2), about 9 months after T1, 1,611 children (51.8% girls) participated again (M = 9.12, SD = 0.93, range 7.12-11.90). At the final wave (T3), about 24 months after T2, 1,501 children (51.7% girls) participated again (M = 11.07, SD = 0.92, range 9.12-13.76). All children who took part at T1 were included in our analyses, and missing data due to dropout or incomplete measures were handled by the Full Information Maximum Likelihood procedure (FIML; Enders & Bandalos, 2001), resulting in a sample size of N = 1,657 children. The children who no longer participated at T3 had significantly higher PHY-A and REL-A, and lower ToM and verbal ability scores at T1 than those children who completed T3.

Study 1

Children were recruited at 33 community primary schools representing a variety of rural and urban areas. About 33.7% of the mothers and 36.9% of the fathers reported holding a university degree, 21.6% and 13.5% a university entrance qualification, 42.9% and 48.0% a vocational level qualification, and 1.8% and 1.7% no or a low level of school graduation.

Material and Procedure

The study was part of a larger research project on intrapersonal developmental risk factors in childhood and adolescence from a longitudinal perspective. Data were collected in individual sessions at the children's schools in a private room by a trained research assistant. Class teachers received individual questionnaires for each participating child. All procedures were approved by the Ethics Committee at the University of Potsdam and by the Ministry of Education, Youth and Sport of the Federal State of Brandenburg, Germany. For each child, informed consent was obtained from the primary caregiver, and the children gave verbal consent.

Theory of Mind. At T1 and T2, cartoons were used to assess ToM (Sebastian et al., 2012). Based on a pilot-study, 12 cartoons (6 for cognitive, 6 for affective ToM) of the original 20 cartoons were chosen as representing an average difficulty level for the studied age group. The cartoons were presented to the children on netbooks using E-Prime 2.0 Professional (Psychology Software Tools, 2012). In each trial, the first three pictures depicted a little story with two characters (A and B). In the affective stories, character A displayed an emotion, for example, sadness after losing a boat on a river (Figure 1); in the cognitive stories, character A displayed an intention or desire, for instance, wanting to pick apples from a tree. Then, two further pictures were presented simultaneously, each displaying a possible ending. The 'correct' ending implied that character B understood A's mental state (e.g., by comforting or helping); in the 'incorrect' ending, character B showed uncaring behavior. The positions of the types of endings were varied across trials, and children were asked to select the picture they thought represented the correct ending by pressing the F-key or the J-key (on a QWERTZ keyboard) for the left or right picture, respectively. Children received feedback only in two initial exercise trials, but not when they worked through the 12 trials. The trials were presented in random order.

Insert Figure 1 about here.

Number of correct responses were summed up separately, resulting in one score for affective ToM and one for cognitive ToM, both with a range of 0-6. In the analyses, we included ToM as a latent factor with the affective and the cognitive ToM score as indicators. Ordinal alpha, which is appropriate for binary items (Gadermann, Guhn, & Zumbo, 2012), indicated good internal consistency for our ToM measure at both time points (α_{T1} = .81, α_{T2} = .82).

Aggressive behavior. At all three time points, aggressive behavior was assessed using 6 items adapted from the Children's Social Behavior Scale - Teacher Form (CSBS-T; Crick, 1996). Classroom teachers indicated how often during the last 6 months the child had shown behaviors of PHY-A (3 items, e.g., "hit, shove, or push peers") and REL-A (3 items, e.g., "try to exclude certain peers from peer group activities"), using a response scale ranging from 1 (*never*), 2 (*once a month or rarely*), 3 (*several times a month*), 4 (*several times a week*), and 5 (*every day)*. In the hypothesis-testing analysis, we used manifest mean scores of PHY-A and REL-A, and included random intercepts for both forms, as described below. Internal consistency was excellent (PHY-A: α_{T1} = .93, α_{T2} = .94, α_{T3} = .93; REL-A: α_{T1} = .91, α_{T2} = .92, α_{T3} = .90).

Verbal ability. Because children's ToM skills can be confounded with their verbal ability (Milligan, Astington, & Dack, 2007), it is essential to control for the latter. At T1, we used a vocabulary test from the Potsdam-Illinois Test for Psycholinguistic Abilities (P-ITPA; Esser, Wyschkon, Ballaschk, & Hänsch, 2010) to assess children's word knowledge and ability to detect relations between words. Scores were normed for children's age.

Study 1

Statistical Analyses

Structural equation models (SEM) were conducted with Mplus (Version 7.4; Muthén & Muthén, 1998-2015), and descriptive analyses were computed with SPSS (Version 23). In all SEMs, the robust maximum-likelihood estimator (MLR) was used to account for deviations from normality of the data. As recommended by Hamaker, Kuiper, and Grasman (2015) we included a latent random intercept for both forms of aggressive behavior to control for stable between-person differences. This procedure provides a more accurate examination of within-person changes. As ToM was measured only at two time points, we were not able to include a random intercept for this construct (random intercepts require at least three time points). Instead, we included ToM as a latent factor with the affective and the cognitive ToM scores as indicators. Metric measurement invariance over time could be established for ToM.

Age and gender were covariates for all constructs, and verbal ability assessed at T1 was an additional covariate for ToM. In addition, separate multi-group analyses were conducted to examine potential gender and age differences. A model in which all paths were restricted to be equal between the groups (fully constrained) was compared with a model in which all paths were freed between the groups (fully unconstrained). Model comparisons were based on χ^2 differences used with scale corrections as proposed by Satorra and Bentler (2001) due to the application of the MLR estimator. Age-groups ('younger' vs. 'middle' vs. 'older') were defined based on splitting the sample by age at T1 at the 33th and 66th percentile ($Md_{33\%}$ = 7.83 years, $Md_{66\%}$ = 8.86 years).

The percentage of missing values ranged from 1.6% to 33.3% (for the sample size of each measure see Table 1). The rate of missing data was highest on the teacher-reports of aggression, due to not-returned questionnaires. The FIML procedure was implemented in all analyses to handle missing data (Enders & Bandalos, 2001).

Overall model fit was evaluated according to Hu and Bentler (1999), with RMSEA ≤ .06, CFI ≥ .95, TLI ≥ .95, and SRMR ≤ .08 indicating a good fit. The χ^2-statistic is reported for providing complete model-fit information, but it was not evaluated as an absolute index of model fit because it is biased for large samples (Schermelleh-Engel, Moosbrugger, & Müller, 2003). Effect sizes for *t*-tests are reported as Cohen's *d* with the conventional interpretation of $d = .20$ as small, $d = .50$ as medium, and $d = .80$ as large effects (Cohen, 1988). For our SEM, we report standardized regression coefficients that can be interpreted in a similar way as Cohen's *d*.

Results

Descriptive Statistics, Gender Differences, and Correlations

Table 1 presents the descriptive statistics and the results of statistical tests for gender differences on the aggressive behavior measures as well as the ToM scores. Teacher-rated PHY-A scores were significantly higher for boys than for girls at each data wave, with medium to large effect sizes (*d*s ≥ 0.66). For REL-A, the gender difference was small (*d*s ≤ 0.14), but boys were rated as significantly more relationally aggressive than girls at T2 and T3. With regard to ToM, no significant gender differences emerged.

PHY-A scores decreased significantly from T1 to T2, $t(1117) = -2.68$, $p < .01$, $d = .06$, but were stable between T2 and T3, $t(843) = -0.99$, $p = .32$, $d = .03$. REL-A scores did not vary significantly between T1 and T2, $t(1115) = 1.86$, $p = .06$, $d = .05$, or between T2 and T3, $t(840) = -1.66$, $p = .10$, $d = .06$. ToM scores increased significantly from T1 to T2, $t(1543) = 16.22$, $p < .001$, $d = .48$.

Insert Table 1 about here.

Bivariate correlations between all variables are displayed in Table 2. Age was positively correlated with ToM at T1 and T2, but uncorrelated with aggressive behavior. All correlations between ToM and aggressive behavior were negative. Stability of ToM between T1 and T2 was relatively low, whereas PHY-A and REL-A showed moderate to high stabilities between the time points. PHY-A and REL-A were strongly positively associated with each other at all time points.

Insert Table 2 about here.

Hypothesis Testing

To test our first set of hypotheses, we computed a SEM with latent factors for ToM at T1 and T2, and manifest mean scores for PHY-A and REL-A at T1, T2, and T3, including random intercepts for both REL-A and PHY-A. In doing so, we investigated the longitudinal reciprocal relations between ToM and PHY-A and REL-A, respectively, controlling for earlier levels of ToM and for stable between-person differences in aggressive behavior. Furthermore, all constructs were controlled for age and gender, and ToM was additionally controlled for T1 verbal ability. Metric measurement invariance over time was established in a step-up approach for PHY-A, REL-A, and ToM from T1 to T2. The overall fit of our final model was good, RMSEA = .03, 90% CI [.02, .04]; CFI = .99; TLI = .98; SRMR = .03; $\chi^2(33) = 71.38$, $p < .001$. The resulting model is presented in Figure 2.

Reciprocal effects of ToM, REL-A, and PHY-A. In line with our hypotheses, we found significant negative paths from ToM at T1 to PHY-A at T2 ($\beta = -.08$, $p < .05$), and from ToM at T2 to PHY-A at T3 ($\beta = -.12$, $p < .01$). Similarly, the hypothesized paths from ToM at T1 to REL-A at T2 ($\beta = -.09$, $p < .05$) as well as from ToM at T2 to REL-A at T3 ($\beta = -.11$, $p < .05$) were significant and negative. The reverse negative path was confirmed only from REL-A at T1 to ToM at T2 ($\beta = -.12$, $p < .05$), but not from PHY-A at T1 to ToM at T2 ($\beta = .09$, $p = .13$). In addition to the expected paths, a negative path from PHY-A at T2 to REL-A at T3 was found ($\beta = -.16$, $p < .01$).

Gender differences. To examine gender differences, we compared the fully constrained multi-group model by gender with the fully unconstrained model (fit constrained model: RMSEA = .05, 90% CI [.04, .05]; CFI = .95; TLI = .93; SRMR = .06; $\chi^2(95) = 256.03$, $p < .001$; fit unconstrained model: RMSEA = .03, 90% CI [.03, .04];

CFI = .98; TLI = .96; SRMR = .04; $\chi^2(65) = 127.14$, $p < .001$). The difference between these two models was significant, $\Delta\chi^2(30) = 119.43$, $p < .001$. In a next step, we computed a revised and final model, in which we constrained all paths to be equal, except the paths that were found to differ significantly between boys and girls in the fully unconstrained model. This revised model showed a good model fit, RMSEA = .03, 90% CI [.02, .04]; CFI = .98; TLI = .97; SRMR = .04; $\chi^2(86) = 153.49$, $p < .001$, which was significantly better than that of the fully constrained model, $\Delta\chi^2(9) = 79.87$, $p < .001$, and not significantly worse than that of the fully unconstrained model, $\Delta\chi^2(21) = 27.83$, $p = .145$. Based on the revised model, there were no significant gender differences in the hypothesized paths (ps ≥ .12). The only significant gender differences were found in the concurrent correlations between PHY-A and REL-A, which were stronger for boys (rs ranging from .54 to .69, ps < .001) than for girls (rs ranging from .42 to .45, ps < .001). In addition, on a within-person level, REL-A at T1 predicted REL-A at T2 only in girls ($\beta = .26$, $p < .001$), but not in boys ($\beta = .09$, $p = .17$).

Age differences. To examine age differences, we followed the same procedure with the three age-groups; fit constrained model: RMSEA = .03, 90% CI [.02, .03]; CFI = .98; TLI = .98; SRMR = .06; $\chi^2(164) = 227.32$, $p < .001$; fit unconstrained model: RMSEA = .03, 90% CI [.02, .04]; CFI = .99; TLI = .97; SRMR = .05; $\chi^2(104) = 154.20$, $p < .001$). The fully constrained model did not fit worse than the fully unconstrained model, $\Delta\chi^2(60) = 75.41$, $p = .09$, indicating that there was no significant moderation by age.

Discussion

The primary aim of this research was to study the reciprocal relations between ToM (assessed by children’s responses to a cartoon task) and teacher-reported PHY-A and REL-A over time in a large community sample of girls and boys in middle childhood. Our longitudinal study with three waves of measurement covered a time period of about 3 years. Past studies on the association of ToM and aggression were mostly cross-sectional, were conducted at preschool age, and did not separate distinct forms of aggression. Therefore, we investigated the reciprocal relations between ToM and PHY-A and REL-A during middle childhood in a prospective design using structural-equation modeling, controlling for age, gender, and verbal ability as well as earlier levels of ToM and partialling out stable between-person differences in aggressive behavior.

Consistent with our predictions, we found that ToM was a significant negative predictor of both PHY-A and REL-A. This was true for the paths from T1 to T2 as well as from T2 to T3. These results indicate that children with lower ToM scores were rated by their teachers to be more physically and relationally aggressive at a later occasion, while controlling for stable between-person differences in aggressive behavior. Even though we did not examine mediating processes in our study, these results are in line with the predictions derived from the SIP model (Crick & Dodge, 1994; Harvey et al., 2001) that deficits in ToM may lead to biased or deficient SIP, which in turn may lead to more aggressive behavior. For example, previous research revealed that the encoding and interpretation of situational cues were less accurate in children with early deficits in ToM (e.g., Choe, Lane, Grabell, & Olson, 2013). Our findings do not support the proposition by Sutton et al. (1999a) that children may use their proficient ToM to manipulate others, joining other studies that also failed to find differences in ToM between bullies and noninvolved children (e.g., Gini, 2006). However, the consideration of potential moderating variables, such as prosocial behavior (e.g., Renouf et al., 2010), may be important to identify those children that exploit their ToM skills in an aggressive fashion.

Although the effect sizes were small, the present findings are consistent with previous studies that examined relations between ToM and aggression (e.g., Kokkinos et al., 2016). Considering the complex nature of the emergence of aggressive behavior as postulated by the SIP model (Crick & Dodge, 1994), many intra- and interpersonal factors can have an impact (e.g., see Krahé, 2013, for a review). Our study demonstrated ToM to be an important, but certainly not the only intrapersonal predictor for the development of aggressive behavior during middle childhood.

Turning to the postulated reverse paths from aggression to ToM, higher scores in REL-A, but not PHY-A, at T1 predicted lower ToM skills at T2. The negative path from REL-A to ToM is in line with the SIP model, postulating reciprocal paths between social cognition and aggressive behavior (Crick & Dodge, 1994). The reason for the missing path from PHY-A to ToM may be that the different functions of aggression in terms of proactive versus reactive aggression were not considered in our study. In past research, only reactive PHY-A was linked to low socio-cognitive abilities (Crick & Dodge, 1996). Proactive PHY-A, however, is an instrumental and planned behavior (Vitaro & Brendgen, 2005) and may be positively linked to ToM. These opposing relations could have cancelled out each other to eliminate the path from PHY-A to ToM. Nevertheless, it remains unclear why this would have been the case only for PHY-A and not for REL-A.

In addition to the hypothesized paths, the analyses revealed a significant negative path from PHY-A at T2 to REL-A at T3. This result indicates that physically aggressive children tend to reduce their relationally aggressive behavior at a later occasion. This is in contrast to previous longitudinal studies that have found a positive relation or no relation between these two forms of aggression over time (e.g., Card, Stucky, Sawalani, & Little, 2008). However, these studies usually examined aggression on a population-level (e.g., by applying traditional cross-lagged panel models), whereas our study examined aggression on the level of the individual (by including random intercepts, Hamaker et al., 2015). The divergence in findings may be an expression of Simpson's paradox (Kievit, Frankenhuis, Waldorp, & Borsboom, 2013), that is, the direction of a link depends on the level of analysis (i.e., population vs. individual). To our knowledge, there are only few studies to date that examined PHY-A and REL-A from a person-centered perspective. These studies examined trajectory groups of PHY-A and REL-A as well as the co-occurrence of these groups (e.g., Côté et al., 2007;

Ettekal, & Ladd, 2015), and found substantial proportions of children who – on an individual level – showed different trajectories of PHY-A and REL-A (e.g., increasing in PHY-A, but decreasing in REL-A). Therefore, future studies should further investigate the interplay of PHY-A and REL-A longitudinally on an intra-personal as well as on an inter-personal level.

In further analyses, we explored potential age and gender differences in the reciprocal relations between ToM, PHY-A, and REL-A in middle childhood. We did not find any differences in the reciprocal relations of ToM and the forms of aggression between girls and boys, and between age-groups. However, we found that the correlation between REL-A and PHY-A was higher for boys than for girls, which is consistent with previous research (Card et al., 2008). In addition, we discovered significant gender differences in the mean level of aggressive behavior. At each wave, boys were rated by their teachers to show significantly more PHY-A than girls. This difference was of medium to large size and is in accordance with earlier studies (e.g., Lansford et al., 2012). Similarly, boys were rated to show more REL-A than girls at T2 and T3. This difference was of small size and is in contrast to other studies that found girls as compared to boys to exhibit more REL-A (e.g., Card et al., 2008; Ostrov et al., 2013); albeit, some studies also found no gender differences in REL-A (e.g., Lansford et al., 2012), or higher scores for boys (e.g., Henington et al., 1998). The present results may have been due to characteristics of our male or female subsample or to population differences, which we could not examine in our study. A more fine-grained analysis of gender differences in REL-A remains an important topic for future research. Regarding age, we did not find differences in the reciprocal relations between ToM and the two forms of aggression across the age-groups of our sample. There were no correlations of age and PHY-A or REL-A, but there were positive correlations of age and ToM. In line with previous research (e.g., Devine & Hughes, 2013), this indicates that ToM is still improving with increasing age in middle childhood. Altogether, we conclude that apart from gender differences in the absolute level of aggressive behavior, in the concurrent correlations between REL-A and PHY-A, and bivariate correlations between age and ToM skills, the prospective relations between aggression and ToM are not moderated by gender or age.

In conclusion, the present findings provided first longitudinal evidence on reciprocal relations between ToM and PHY-A, as well as REL-A, as distinct forms of aggression, in middle childhood. Our results of negative reciprocal relations between ToM and REL-A or PHY-A, respectively, are consistent with the 'social skills deficit-view' of Crick and Doge (1994), and they do not support the proposition by Sutton et al. (1999a) that aggressive children have advanced ToM abilities. This suggests that improving ToM abilities may reduce rather than promote PHY-A and REL-A in middle childhood. The impact of REL-A on children's well-being is usually evaluated by teachers and parents as being less severe than that of PHY-A (e.g., Hurd & Gettinger, 2011). In contrast, our study showed that trivialization of REL-A may have detrimental effects on socio-cognitive development. Therefore, interventions that aim at reducing REL-A and/or fostering ToM would be indicated for interrupting the vicious circle of deficient socio-cognitive abilities promoting aggressive behavior and vice versa. Following our results, gender differences do not seem to play a prominent role in the examined processes. Therefore, interventions should equally focus on girls and boys.

Strengths and Limitations

We believe our study has several strengths. It is based on a large sample of children attending community elementary schools, and it includes three data waves covering middle childhood over a total of about 3 years. We controlled for stable between-person differences in aggression and for past levels of ToM (construct stability), fostering a causal interpretation of the paths (Marmor & Montemayor, 1977). Further, we distinguished between two forms of aggression, PHY-A and REL-A, which have been found in some previous studies to differ between girls and boys (e.g., Lansford et al., 2012), and we explored potential gender and age differences. Finally, as one of the first studies, we investigated not only the path from ToM to the two forms of aggression, but also the reverse paths. Crick and Dodge (1994) already emphasized the importance of studying the effect of social experiences, such as aggressive behavior, on social cognitions. This can promote the understanding of how children develop outcome expectations for social situations, which they then apply in reaching response decisions.

At the same time, there are some limitations to our study. First, we used the prominent SIP model by Crick and Dodge (1994) as a framework for our study, but we did not include measures specifically designed for SIP, for example video vignettes (Lansford et al., 2006). Consequently, the specific underlying mechanisms by which ToM affects SIP cannot be inferred. Further, we did not examine other mediating mechanisms, such as empathy or moral disengagement (e.g., Kokkinos et al., 2016). Future studies should therefore integrate measures for specific cognitive processes that are assumed to influence aggressive behavior.

Second, the time intervals differed between T1 and T2 (about 9 months), and T2 and T3 (about 24 months). Consequently, prospective effects were compared between unequal time periods; however, it is even more remarkable that ToM predicted PHY-A and REL-A at both time intervals. Another limitation is that our ToM measure appeared somewhat too easy for our sample, especially at T2. These ceiling effects reduced the variance in the ToM measures, making it harder to detect potential effects and to identify children that are very skilled in ToM. Future studies should continue to develop time- and resource-efficient ToM measures that create sufficient variability in middle childhood. Finally, the use of teacher-reports to assess aggressive behavior can be considered as a limitation of our study. Teachers may not be fully aware of the extent to which children engage in aggressive behavior, in particular when it comes to REL-A, which includes more covert behavior (e.g., spreading rumors).

Despite these limitations, our findings highlight the consideration of reciprocal relations between ToM and aggressive behavior over the course of middle childhood. Future research should further investigate how aggressive behavior and ToM develop over the whole course of childhood and how this feeds back into the development of the respective other part over time. On the long run, this information can help to reduce the detrimental effects of aggressive behavior on children's lives and health.

Study 1

References

Archer, J., & Coyne, S. (2005). An integrated review of indirect, relational, and social aggression. *Personality and Social Psychology Review*, *9*, 212-230. doi:10.1207/s15327957pspr0903_2

Baron, R. A., & Richardson, D. S. (1994). *Human aggression* (2nd ed.). New York, NY: Plenum Press.

Capage, L., & Watson, A. C. (2001). Individual differences in theory of mind, aggressive behavior, and social skills in young children. *Early Education & Development*, *12*, 613–628. doi:10.1207/s15566935eed1204_7

Card, N., Stucky, B., Sawalani, G., & Little, T. (2008). Direct and indirect aggression during childhood and adolescence: A meta-analytic review of gender differences, intercorrelations, and relations to maladjustment. *Child Development*, *79*, 1185–1229. doi: 10.1111/j.1467-8624.2008.01184.x

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*, 2245–2256. doi:10.1037/a0032293

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Côté, S. M., Vaillancourt, T., Barker, E. D., Nagin, D., & Tremblay, R. E. (2007). The joint development of physical and indirect aggression: Predictors of continuity and change during childhood. *Development and Psychopathology*, *19*, 37–55. doi:10.1017/S0954579407070034

Crick, N. R., Casas, J. F., & Ku, H.-C. (1999). Relational and physical forms of peer victimization in preschool. *Developmental Psychology*, *35*, 376–385. doi:10.1037/0012-

1649.35.2.376

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin*, *115*, 74–101. doi:10.1037/0033-2909.115.1.74

Crick, N. R., & Dodge, K. A. (1996). Social information-processing mechanisms in reactive and proactive aggression. *Child Development*, *67*, 993–1002. doi:10.1111/j.1467-8624.1996.tb01778.x

Crick, N. R., & Grotpeter, J. K. (1995). Relational aggression, gender, and social-psychological adjustment. *Child Development*, *66*, 710–722. doi:10.1111/j.1467-8624.1995.tb00900.x

Crick, N. R., Nelson, D. A., Morales, J. R., Cullerton-Sen, C., Casas, J. F., & Hickman, S. E. (2001). Relational victimization in childhood and adolescence: I hurt you through the grapevine. In J. Juvonen & S. Graham (Eds.), *Peer harassment in school: The plight of the vulnerable and victimized* (pp. 196–214). New York, NY: Guilford Press. doi:0009-3920/2002/7304-0010

Crick, N. R., Ostrov, J. M., Burr, J. E., Cullerton-Sen, C., Jansen-Yeh, E., & Ralston, P. (2006). A longitudinal study of relational and physical aggression in preschool. *Journal of Applied Developmental Psychology*, *27*, 254–268. doi:10.1016/j.appdev.2006.02.006

Dahlberg, L. L., & Krug, E. G. (2002). Violence - a global public health problem. In E. G. Krug, L. L. Dahlberg, J. A. Mercy, A. B. Zwi, & R. Lozano (Eds.), *World report on violence and health* (pp. 3–21). Geneva, CH: World Health Organisation.

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*, 105–116. doi:10.1207/s15374424jccp3401

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*, 916–934. doi:10.1111/1467-8624.00447

Devine, R. T., & Hughes, C. (2013). Silent films and strange stories: Theory of mind, gender, and social experiences in middle childhood. *Child Development*, *84*, 989–1003. doi:10.1111/cdev.12017

Enders, C., & Bandalos, D. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling, 8*, 430–457. doi:10.1207/S15328007SEM0803_5

Esser, G., Wyschkon, A., Ballaschk, K., & Hänsch, S. (2010). *P-ITPA: Potsdam-Illinois Test für Psycholinguistische Fähigkeiten* [P-ITPA: Potsdam-Illinois test for psycholinguistic abilities]. Göttingen, Germany: Hogrefe.

Ettekal, I., & Ladd, G. W. (2015). Costs and benefits of children's physical and relational aggression trajectories on peer rejection, acceptance, and friendships: Variations by aggression subtypes, gender, and age. *Developmental Psychology*, *51*, 1756–1770. doi:10.1037/dev0000057

Fahie, C. M., & Symons, D. K. (2003). Executive functioning and theory of mind in children clinically referred for attention and behavior problems. *Journal of Applied Developmental Psychology*, *24*, 51–73. doi:10.1016/S0193-3973(03)00024-8

Frith, U., & Frith, C. (2010). The social brain: Allowing humans to boldly go where no other species has been. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *365*, 165–176. doi:10.1098/rstb.2009.0160

Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation*, *17*(3), 1–13.

Gini, G. (2006). Social cognition and moral cognition in bullying: what's wrong? *Aggressive Behavior*, *32*, 528-539. doi: 10.1002/ab.20153

Study 1

Gomez-Garibello, C., & Talwar, V. (2015). Can you read my mind? Age as a moderator in the relationship between theory of mind and relational aggression. *International Journal of Behavioral Development*, *39*, 552–559. doi:10.1177/0165025415580805

Grueneisen, S., Wyman, E., & Tomasello, M. (2015). "I know you don't know I know..." Children use second-order false-belief reasoning for peer coordination. *Child Development*, *86*, 287–293. doi:10.1111/cdev.12264

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*, 102–116. doi:10.1037/a0038889

Harvey, R. J., Fletcher, J., & French, D. J. (2001). Social reasoning: An influence on aggression. *Clinical Psychology Review*, *21*, 459–469. doi: 10.1016/S0272-7358(99)00068-9

Henington, C., Hughes, J. N., Cavell, T. A., & Thompson, B. (1998). The role of relational aggression in identifying aggressive boys and girls. *Journal of School Psychology, 36*, 457–477. doi:10.1016/S0022-4405(98)00015-6

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. doi:10.1080/10705519909540118

Huesmann, R. L. (1998). The role of social information processing and cognitive schema in the acquisition and maintenance of habitual aggressive behavior. In R. G. Geen & E. Donnerstein (Eds.), *Human aggression. Theories, research, and implications for social policy* (pp. 73–109). San Diego, CA: Academic Press.

Hughes, C., & Dunn, J. (1998). Understanding mind and emotion: Longitudinal associations with mental-state talk between young friends. *Developmental Psychology*, *34*, 1026–1037. doi:10.1037/0012-1649.34.5.1026

Hughes, C., & Devine, R. T. (2015). Individual differences in theory of mind from preschool to adolescence: Achievements and directions. *Child Development Perspectives, 9*, 149-

153. doi:10.1111/cdep.12124

Hughes, C., & Ensor, R. (2006). Behavioral problems in 2-year-olds: Links with individual differences in theory of mind, executive function and harsh parenting. *Journal of Child Psychology and Psychiatry*, *47*, 488–497. doi:10.1111/j.1469-7610.2005.01519.x

Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social Development*, *13*, 590–619. doi:10.1111/j.1467-9507.2004.00285.x

Hurd, H. D., & Gettinger, M. (2011). Mothers' and teachers' perceptions of relational and physical aggression in pre-school children. *Early Child Development and Care*, *181*, 1343–1359. doi:10.1080/03004430.2010.527336

Kievit, R. A., Frankenhuis, W. E., Waldorp, L. J., & Borsboom, D. (2013). Simpson's paradox in psychological science: A practical guide. *Frontiers in Psychology*, *4*, Article 513. doi:10.3389/fpsyg.2013.00513

Kokkinos, C. M., Voulgaridou, I., Mandrali, M., & Parousidou, C. (2016). Interactive links between relational aggression, theory of mind, and moral disengagement among early adolescents. *Psychology in the Schools*, *53*, 253–269. doi:10.1002/pits.21902

Krahé, B. (2013). *The social psychology of aggression* (2nd ed.). Hove, UK: Psychology Press.

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*, 715–724. doi: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*, 593–602. doi:10.1017/S0954579410000301

Study 1

Lansford, J. E., Skinner, A. T., Sorbring, E., Di Giunta, L., Deater-Deckard, K., Dodge, K. A., … Chang, L. (2012). Boys' and girls' relational and physical aggression in nine countries. *Aggressive Behavior*, *38*, 298–308. doi:10.1002/ab.21433

Marmor, G. S., & Montemayor, R. (1977). The cross-lagged panel design: A review. *Perceptual and Motor Skills, 45*, 883–893. doi:10.2466/pms.1977.45.3.883

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, *78*, 622–646. doi:10.1111/j.1467-8624.2007.01018.x

Murray-Close, D., Ostrov, J. M., & Crick, N. R. (2007). A short-term longitudinal study of growth of relational aggression during middle childhood: Associations with gender, friendship intimacy, and internalizing problems. *Development and Psychopathology*, *19*, 187–203. doi:10.1017/S0954579407070101

Muthén, L. K., & Muthén, B. O. (1998-2015). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Olweus, D. (2013). School bullying: Development and some important challenges. *Annual Review of Clinical Psychology*, *9*, 751–780. doi:10.1146/annurev-clinpsy-050212-185516

Ostrov, J. M., Murray-Close, D., Godleski, S. A., & Hart, E. J. (2013). Prospective associations between forms and functions of aggression and social and affective processes during early childhood. *Journal of Experimental Child Psychology*, *116*, 19-36. doi:10.1016/j.jecp.2012.12.009

Psychology Software Tools. (2012). E-Prime 2.0 [Computer Software]. Retrieved from http://www.pstnet.com

Renouf, A., Brendgen, M., Parent, S., Vitaro, F., Zelazo, P. D., Boivin, M., ... Séguin, J. R. (2010). Relations between theory of mind and indirect and physical aggression in kindergarten: Evidence of the moderating role of prosocial behaviors. *Social*

Development, *19*, 535–555. doi:10.1111/j.1467-9507.2009.00552.x

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514. doi:10.1007/BF02296192

Schermelleh-Engel, H., Moosbrugger, K., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*, 23-74.

Sebastian, C. L., Fontaine, N. M. G., Bird, G., Blakemore, S.-J., De Brito, S. A., McCrory, E. J. P., & Viding, E. (2012). Neural processing associated with cognitive and affective theory of mind in adolescents and adults. *Social Cognitive and Affective Neuroscience*, *7*, 53–63. doi:10.1093/scan/nsr023

Shamay-Tsoory, S. G., Harari, H., Aharon-Peretz, J., & Levkovitz, Y. (2010). The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. *Cortex*, *46*, 668–677. doi:10.1016/j.cortex.2009.04.008

Song, J.-H., Volling, B. L., Lane, J. D., & Wellman, H. M. (2016). Aggression, sibling antagonism, and theory of mind during the first year of siblinghood: A developmental cascade model. *Child Development*, *87*, 1250–1263. doi:10.1111/cdev.12530

Sutton, J., Smith, P. K., & Swettenham, J. (1999a). Bullying and “theory of mind”: A critique of the “social skills deficit” view of anti-social behavior. *Social Development*, *8*, 117–127. doi:10.11111467-9507.00083

Sutton, J., Smith, P. K., & Swettenham, J. (1999b). Social cognition and bullying: Social inadequacy or skilled manipulation? *British Journal of Developmental Psychology*, *17*, 435-450. doi:10.1348/026151099165384

Teufel, C., Fletcher, P. C., & Davis, G. (2010). Seeing other minds: Attributed mental states influence perception. *Trends in Cognitive Sciences*, *14*, 376–382. doi:10.1016/j.tics.2010.05.005

Vetter, N., Weigelt, S., Döhnel, K., Smolka, M. N., Kliegel, M. (2014). Ongoing neural development of affective theory of mind in adolescence. *Social Cognitive and Affective Neuroscience*, *9*, 1022–1029. doi:10.1093/scan/nst081

Vitaro, F., & Brendgen, M. (2005). Proactive and reactive aggression: A developmental perspective. In R. E. Tremblay, W. W. Hartup, & J. Archer (Eds.), *Developmental origins of aggression* (pp. 178–201). New York, NY: Guilford Press.

Study 1

Table 1

Descriptive Statistics and Gender Differences

	n (*n girls*)	Range	Total *M* (*SD*)	Girls *M* (*SD*)	Boys *M* (*SD*)	Gender difference
Physical Aggression[a]						
T1	1411	1 - 5	1.49	1.22	1.79	$t(1063.36) = -14.52$***
	(732)		(0.79)	(0.52)	(0.91)	$d = 0.77$
T2	1148	1 - 5	1.48	1.18	1.79	$t(825.45) = -14.23$***
	(585)		(0.77)	(0.46)	(0.90)	$d = 0.85$
T3	1107	1 - 5	1.44	1.20	1.67	$t(864.37) = -11.07$***
	(551)		(0.75)	(0.49)	(0.88)	$d = 0.66$
Relational Aggression[a]						
T1	1408	1 - 5	1.50	1.47	1.54	$t(1406) = -1.65$
	(732)		(0.70)	(0.70)	(0.70)	$d = 0.10$
T2	1147	1 - 5	1.56	1.51	1.60	$t(1145) = -2.15$*
	(584)		(0.76)	(0.73)	(0.79)	$d = 0.12$
T3	1105	1 - 5	1.50	1.45	1.55	$t(1086.01) = -2.22$*
	(550)		(0.71)	(0.66)	(0.76)	$d = 0.14$
Theory of Mind[b]						
T1	1630	0 - 12	10.09	10.17	10.00	$t(1628) = 1.83$
	(848)		(1.90)	(1.86)	(1.95)	$d = 0.09$
T2	1567	0 - 12	10.91	10.92	10.89	$t(1565) = 0.41$
	(800)		(1.44)	(1.46)	(1.42)	$d = 0.02$

Notes. [a] mean score; [b] sum score.

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

Study 1

Table 2

Intercorrelations of Study Variables

	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. T1 Age	-.11***	.16***	-.03	-.05	.15***	-.03	-.04	-.02	-.02
2. T1 Verbal Ability		.15***	-.09**	-.09***	.17***	-.10**	-.11***	-.13***	-.14***
3. T1 Theory of Mind			-.12***	-.05	**.33*****	-.11***	-.08**	-.11***	-.12***
4. T1 Physical Aggression				.62***	-.07**	**.72*****	.45***	**.48*****	.32***
5. T1 Relational Aggression					-.09***	.45***	**.53*****	.33***	**.36*****
6. T2 Theory of Mind						-.08**	-.08**	-.08**	-.08*
7. T2 Physical Aggression							.66***	**.50*****	.37***
8. T2 Relational Aggression								.31***	**.40*****
9. T3 Physical Aggression									.66***
10. T3 Relational Aggression									

Notes. Stabilities are highlighted in bold; *N* = 1,657.

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

Study 1

Figure Captions

Figure 1. Example Cartoon Story 'Boat' for Assessing Affective ToM.

Notes. The first three pictures were shown consecutively; after that, two pictures displaying possible endings appeared simultaneously (here, the correct answer has a bold frame).

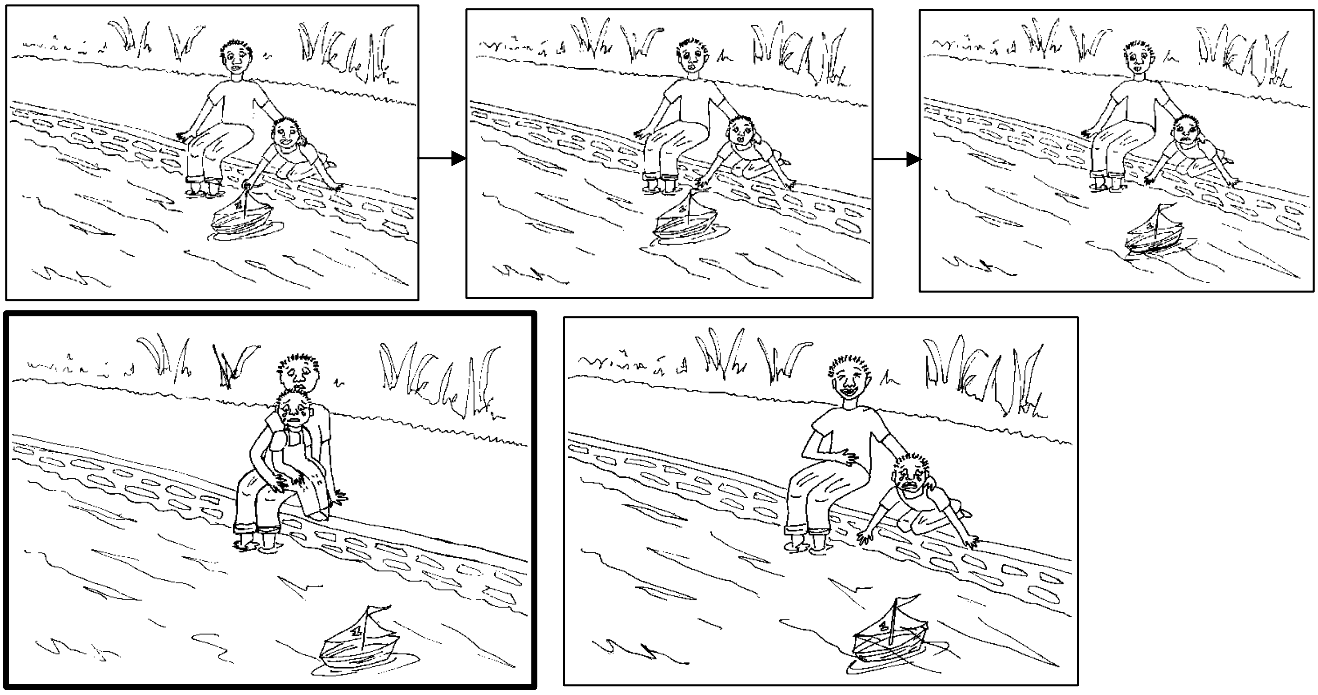
Figure 2. SEM for Physical Aggression, Relational Aggression and Theory of Mind from T1 to T3.

Notes. Standardized coefficients are displayed; dashed lines display non-significant paths; numbers in square brackets display 95% CIs; mean relational aggression and mean physical aggression partial out between-person stability over time (random intercept); all variables were controlled for age and gender, theory of mind was additionally controlled for verbal ability; $N = 1{,}657$; Model fit: RMSEA = .03, 90% CI [.02, .04]; CFI = .99; TLI = .98; SRMR = .03; $\chi^2(33) = 71.38$.

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

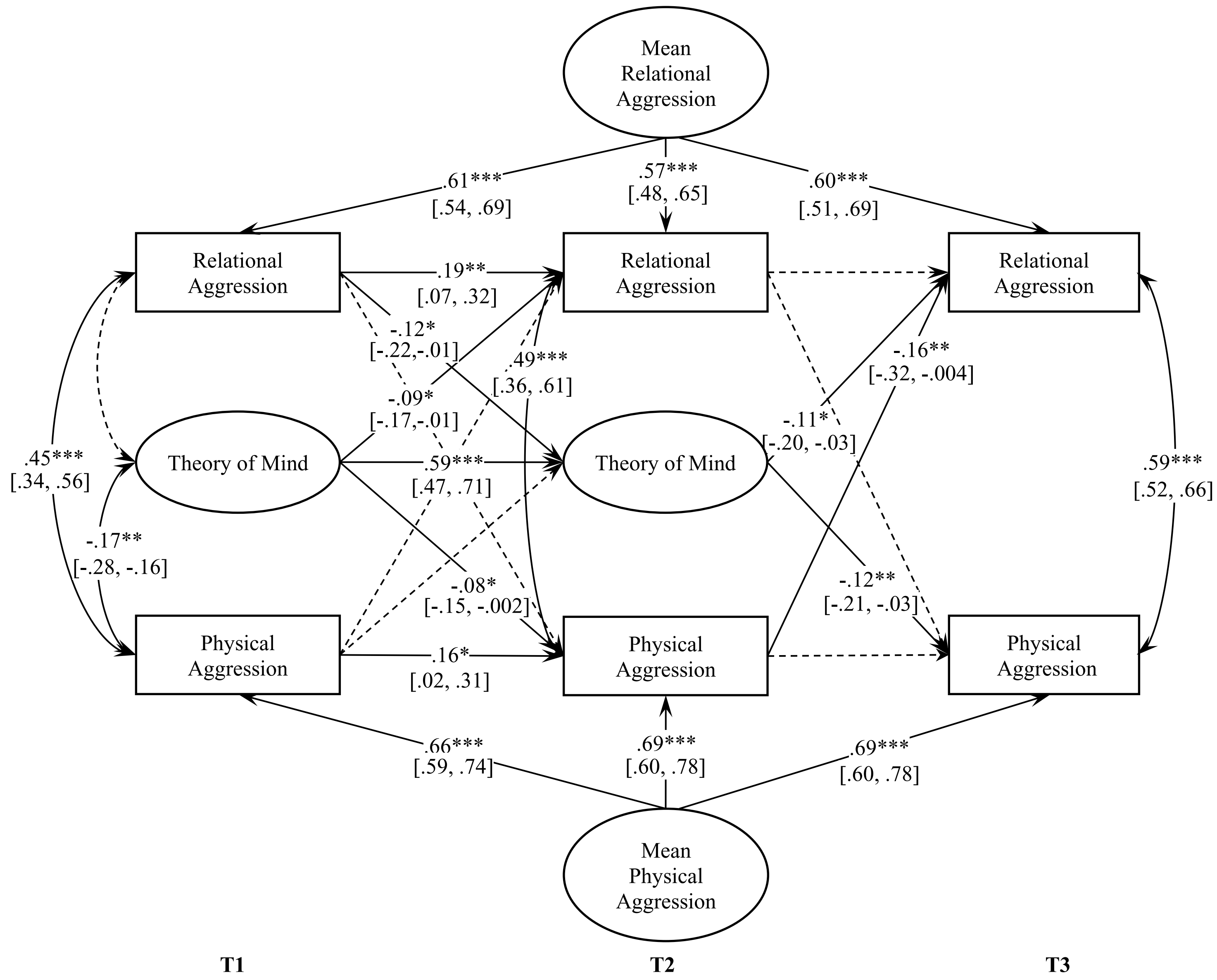
Study 1

Figure 1.



Study 1

Figure 2.



4. Study 2

Rohlf, H., Holl, A. K., Kirsch, F., Krahé, B., & Elsner, B. (2018). Longitudinal Links between Executive Function, Anger, and Aggression in Middle Childhood. *Frontiers in Behavioral Neuroscience*, *12*(27).
http://doi.org/10.3389/fnbeh.2018.00027

Study 2

Abstract

Previous research has indicated that executive function (EF) is negatively associated with aggressive behavior in childhood. However, there is a lack of longitudinal studies that have examined the effect of deficits in EF on aggression over time and taken into account different forms and functions of aggression at the same time. Furthermore, only few studies have analyzed the role of underlying variables that may explain the association between EF and aggression. The present study examined the prospective paths between EF and different forms (physical and relational) and functions (reactive and proactive) of aggression. The habitual experience of anger was examined as a potential underlying mechanism of the link between EF and aggression, because the tendency to get angry easily has been found to be both a consequence of deficits in EF and a predictor of aggression. The study included 1,652 children (between 6 and 11 years old at the first time point), who were followed over three time points (T1, T2, and T3) covering 3 years. At T1, a latent factor of EF comprised measures of planning, rated via teacher reports, as well as inhibition, set shifting, and working-memory updating, assessed experimentally. Habitual anger experience was assessed via parent reports at T1 and T2. The forms and functions of aggression were measured via teacher reports at all three time points. Structural equation modeling revealed that EF at T1 predicted physical, relational, and reactive aggression at T3, but was unrelated to proactive aggression at T3. Furthermore, EF at T1 was indirectly linked to physical aggression at T3, mediated through habitual anger experience at T2. The results indicate that deficits in EF influence the later occurrence of aggression in middle childhood, and the tendency to get angry easily mediates this relation.

Study 2

Introduction

A meta-analysis that included a wide range of EF measures concluded that EF is negatively associated to antisocial behavior, with varying effect sizes depending on the specific form of antisocial behavior and the occurrence of comorbid problems (Ogilvie et al., 2011). However, there is a lack of longitudinal studies that examined the effect of deficits in EF on the development of aggression, particularly in middle childhood, taking into account different forms and functions of aggression. The present study extends previous research by examining the longitudinal links between EF and different forms (relational and physical) and functions (reactive and proactive) of aggression over 3 years. In addition, previous research has mostly studied direct links between EF and aggression without considering potential underlying mechanisms. The present study addressed this issue by including individual differences in the experience of anger as a mediating variable.

Executive Function

There is disagreement in the literature over the exact definition of EF. However, EF can be described as an umbrella term that is usually equated with conscious, higher order processes associated with the prefrontal cortex (Hughes et al., 2005). EF governs goal-directed action and planning of behavior, and allows for adaptive responses to novel, complex, or ambiguous situations. As an important aspect of self-regulation, EF is considered vital for autonomous and adaptive psychological functioning (Séguin et al., 2007). Miyake et al. (2000) differentiated between three components of EF in college students, namely inhibition of prepotent responses, working memory updating, and mental set shifting. In a latent-variable analysis, these factors were moderately correlated, but clearly separable, and also had some common underlying mechanisms that contributed to all EF tasks. This unity-but-diversity framework is the most accepted conceptualization of EF, supported also by studies with children and adolescents (e.g., Lehto et al., 2003; Huizinga et al., 2006). Inhibition involves withholding or restraint of a motor response, and is considered central to EF (Miyake et al., 2000). Working memory updating (working memory) is the ability to maintain and manipulate information over brief periods of time (Huizinga et al., 2006). Shifting is the ability to alternate between mental rule sets or tasks (Miyake et al., 2000), and is considered the most complex EF component. An additional EF component that is also frequently mentioned is planning, which is essential to the EF domains of goal setting and goal-oriented behavior (Anderson, 2002). Unlike other cognitive abilities, EF shows a pronounced development after early childhood, paralleling the protracted maturation of the prefrontal cortex (Blakemore and Choudhury, 2006). The single components of EF, however, seem to follow differential courses throughout childhood and adolescence, involving progressive and regressive phases of development (Best et al., 2009; Best and Miller, 2010).

Aggression

Among the many advantages of EF is the ability to regulate behavior that is prohibited by social norms, such as aggressive behavior. Aggression is defined as “any form of behavior directed toward the goal of harming or injuring another living being that is motivated to avoid such treatment” (Baron & Richardson, 1994, p. 7). In the present study, we distinguished between different forms and functions of aggression. A widely used classification of forms of aggressive behavior is the distinction between physical and relational aggression. *Physical aggression* refers to behavior that is intended to harm another person through the threat or use of physical force, whereas *relational aggression* is defined as behavior aimed at damaging another person’s social relationships or feeling of social inclusion (Crick & Grotpeter, 1995). Children’s use of physical aggression normally decreases during their preschool years, whereas relational aggression tends to increase during middle childhood, particularly in girls (Côté et al., 2007).

Study 2

The distinction between different functions refers to the motivation that leads a person to act aggressively. Unprovoked aggressive behavior that aims to reach a certain goal, such as social dominance or the achievement of material goals, is described as *proactive aggression*. Proactively aggressive behavior can also be described as “offensive”, “instrumental”, and “cold-blooded” (Vitaro, Brendgen, & Barker, 2006). Proactive aggression is conceptually linked to callous-unemotional (CU) traits. Children high on CU traits are characterized by a lack of guilt, reduced empathy, reduced display of emotions, callousness and uncaring behavior (Blair, Leibenluft, & Pine, 2014; Vitaro et al., 2006), and they use aggressive behavior to reach desired rewards or social dominance (Pardini, Lochman, & Frick, 2003). *Reactive aggression*, by contrast, refers to aggressive behavior that is displayed in response to a perceived threat or provocation (Card & Little, 2006; Dodge & Coie, 1987). Reactively aggressive behavior can also be described as “defensive”, “impulsive”, and “hot-blooded” (Walters, 2005). The majority of children seem to follow a stable-low course in reactive and proactive aggression over the course of middle childhood, but some children show substantial changes by either increasing or decreasing their use of reactive and/or proactive aggression (Cui et al., 2016). Taken together, longitudinal evidence suggests that middle childhood is a period of important developmental change for both the forms and functions of aggression.

The Link between Executive Function and Aggression

A large body of correlational research has shown that EF is negatively related to aggression in preschool-aged children and adolescents. For instance, low levels of EF coincide with preschoolers’ externalizing behavior, which includes aggressive behavior (e.g., Hughes & Ensor, 2008). Similarly, preschoolers who were rated as “hard to manage” by their parents showed significantly lower EF than a less problematic comparison group (Hughes, Dunn, & White, 1998). A meta-analysis that covered a broad age range from early childhood to adulthood concluded the negative relation between EF and antisocial behavior to be robust, with one of the largest effects for externalizing behavior disorder (Ogilvie et al., 2011). Similarly, in a meta-analysis on preschoolers, EF, inhibition in particular, was correlated with externalizing behavior with a medium effect size (Schoemaker, Mulder, Deković, & Matthys, 2013). Longitudinal evidence demonstrated that 3-year old children with low levels of effortful control, a cognitive construct closely related to EF (Bridgett et al., 2013), showed an increased risk for a chronic pattern of elevated externalizing behavior throughout middle childhood (Olson et al., 2017). By comparison, research on the relation between EF and aggression in middle childhood is limited, although this age range would be important to investigate. As a consequence of the developmental change in the forms (Côté et al., 2007) and functions (Cui et al., 2016) of aggression in middle childhood and the ongoing development of EF (e.g., Blakemore & Choudhury, 2006), this developmental period is of particular interest.

Different theoretical explanations have been proposed for the link between EF and aggression. The frontal-lobe hypothesis of emotional and behavioral regulation suggests that cognitive-neuropsychological functions in the frontal lobe, which are related to EF, appear to be systematically impaired in individuals showing physical aggression (Séguin, 2009). That is, children whose physically aggressive behavior does not decline after preschool as expected (Côté et al., 2007) are thought to have deficits in their EF. As a consequence, they have problems in regulating their behavior and solving social problems (Séguin & Zelazo, 2005; Zadeh et al., 2007). For example, they may not represent a problem adequately, may show deficits in planning a solution, or in reacting flexibly to different kinds of social situations. This is also the focus of another theoretical framework, the Social Information Processing (SIP) model, which proposes that children who show aggressive behavior may have deficits in their social information processing compared to nonaggressive age mates (Crick & Dodge, 1994). These difficulties may be influenced by children's EF (De Castro & van Dijk, 2018; Huesmann, Eron, & Yarmel, 1987).

A further explanation relates to the integration of emotional processes into social-cognitive information processing (Lemerise & Arsenio, 2000). Particularly, anger seems to play an important part in the mediation between social-cognitive processes and aggressive responses (De Castro, Merk, Koops, Veerman, & Bosch, 2005). EF is involved in the regulation of negative affect already in toddlers (Putnam et al., 2008) and preschool children (Carlson & Wang, 2007). Accordingly, deficits in EF may increase the experience of anger. Indeed, deficits in EF are linked to higher levels of negative affect, including anger (Bridgett et al., 2013; Gagne & Hill Goldsmith, 2011; Healey, Marks, & Halperin, 2011). In addition, higher cognitive control – a related construct to EF – in adolescence seems to act as a buffer against later, maladaptive outcomes of chronic anger, for example, adult antisocial personality traits (Hawes et al., 2016). Further evidence comes from research on irritability, which is defined as an increased proneness to anger (Leibenluft, 2017). That is, children with higher levels of irritability showed deficits in the processing of emotional stimuli, impaired context-sensitive regulation (Leibenluft & Stoddard, 2013), and neural dysfunction in processes associated to EF, such as error monitoring, reward processing, and emotion regulation (Perlman et al., 2015). Anger, in turn, is an important impelling factor of aggressive behavior (Leibenluft & Stoddard, 2013). The role of anger as an antecedent of aggression can be explained by the anger-related action tendency that is assumed to activate aggression-related motor impulses (Berkowitz & Harmon-Jones, 2004). Further, irritability can be conceptualized as a maladaptive response to frustration or threat (Leibenluft, 2011). Supporting this assumption, children in preschool age and in middle childhood who are prone to anger were found to be more likely to engage in aggressive behavior (e.g., Arsenio, Cooperman, & Lover, 2000; Eisenberg et al., 1999; Olson, Sameroff, Kerr, Lopez, & Wellman, 2005; Wakschlag et al., 2015). The theoretical and empirical links between EF and anger on the one hand, and anger and aggression on the other hand suggest that the association of EF with aggression may partly be explained by the habitual experience of anger. So far, this assumption has received little attention, especially for the developmental period of middle childhood, and was therefore addressed in the present study.

With regard to different forms of aggression, most research has focused on relations of EF and physical rather than relational aggression. One reason for this may be that – as suggested by the frontal-lobe hypothesis of emotional and behavioral regulation outlined above – impairments in EF appear to be specific to physically aggressive behavior (Séguin, 2009). Those studies that included both forms of aggression revealed mixed findings. In line with the frontal-lobe hypothesis, EF was found to be negatively associated with physical aggression and not related with relational aggression in 3- to 6-year-old children (O'Toole, Monks, & Tsermentseli, 2017). By contrast, other cross-sectional research has found negative relations of EF to both physical and relational aggression in early childhood; only working memory was positively associated with proactive relational aggression (Poland et al., 2016). However, these studies did not account for overlapping variance of physical and relational aggression, which may have an impact on the respective relations. With regard to middle childhood, previous research has failed to support the assumption that only physical aggression is related to deficits in EF. In a sample of fourth- and fifth-grade children, impaired central executive working memory, an indicator of EF, was associated with both physical and relational aggression (McQuade et al., 2013). Furthermore, in a population sample of 9-year-olds, working memory updating was negatively related only to relational, not to physical aggression (Granvald & Marciszko, 2016). Another study in middle childhood did not find significant paths of impaired EF to physical or relational aggression after controlling for symptoms of attention deficit/hyperactivity disorder (Diamantopoulou, Rydell, Thorell, & Bohlin, 2007). Altogether, the inconsistency among studies that have taken both relational and physical aggression into account points to the need for further research into the role of EF in the development of different forms of aggression.

Study 2

Regarding functions of aggression, few studies have differentiated between reactive and proactive aggression when examining the link between EF and aggression, particularly in middle childhood. In 9- to 12-year-old children, deficits in EF, particularly in response inhibition and planning, were found to be positively associated with reactive aggression. The relations between planning and reactive aggression, but not between planning and proactive aggression were moderated by hostile attributional biases (Ellis, Weiss, & Lochman, 2009; Rathert, Fite, & Gaertner, 2011). In addition, a measure of self-regulation that included EF-components was negatively linked to reactive, but not proactive aggression in 6- to 16-year-old children and adolescents (White et al., 2013). Thus, deficits in EF seem to be more involved in the development of reactive compared to proactive aggression. One explanation for the relation between deficits in EF and reactive aggression may be the potential mediating role of anger, as outlined above. Because anger is a major component of reactive, but not proactive aggression (for a review, see Hubbard, Romano, McAuliffe, & Morrow, 2010), it can be assumed that only reactive aggression is indirectly predicted by poor EF via the experience of anger. Consequently, anger may also mediate between EF and reactive aggression.

The Current Study

The aim of this study was to examine the prospective paths between EF and different forms (physical and relational) and functions (reactive and proactive) of aggression in a large population-based sample in middle childhood, with the habitual experience of anger considered as a potential underlying mechanism. The study included three measurement time-points covering three years. At T1, a latent factor of EF was calculated from measures of inhibition, set shifting, working-memory updating, and planning, which were assessed by using behavioral tasks and a teacher-report measure. Children's tendency to experience anger was assessed via parent reports at T1 and T2, and the forms and functions of aggression were rated by teachers at all three time points. The prospective paths were analyzed via structural-equation modeling, controlling for age, gender, and information-processing capacity.

Based on the theoretical assumptions and previous evidence outlined above, four hypotheses were postulated: First, we expected to find a negative relation between EF at T1 and physical aggression at T3, such that lower EF would predict higher levels of later physical aggression (Hypothesis 1). Considering relational aggression, the existing evidence is mixed, because some research found a negative (McQuade et al., 2013) and other either no (Diamantopoulou et al., 2007) or even a positive relation to EF (Poland et al., 2016). We therefore examined these competing predictions for the relation between EF at T1 and relational aggression at T3 in our model. In addition to potential direct effects, we expected negative indirect effects between EF at T1 and physical and relational aggression at T3 through habitual anger at T2 (Hypothesis 2). Thus, we proposed that lower EF would predict a higher tendency to experience anger at T2, which in turn would predict higher rates of physical and relational aggression at T3. With regard to the functions of aggression, we postulated that EF at T1 would be a negative predictor of reactive aggression at T3 but would be unrelated to proactive aggression at T3 (Hypothesis 3), based on earlier evidence (e.g., Ellis et al., 2009; Rathert et al., 2011). Furthermore, we expected negative indirect effects between EF at T1 and reactive aggression at T3 through habitual anger at T2. Thus, lower EF at T1 would predict a higher tendency to experience anger at T2, which in turn would predict higher rates of reactive aggression at T3 (Hypothesis 4).

In addition, we tested potential gender differences in the postulated paths between EF, anger, and aggression. In previous research, gender differences received little attention, in particular in conjunction with the distinction between forms and functions of aggression. Given the gender-related differences in the occurrence of the two forms of aggression (boys usually show more physical, girls slightly more relational aggression; e.g., Card, Stucky, Sawalani, & Little, 2008), it was deemed important to include gender as a potential moderator in the analyses. However, the few studies that have addressed gender differences yielded little support for the assumption that the longitudinal links between EF and forms or functions of aggression might differ by gender (White et al., 2013). Based on these findings, we expected that the proposed associations would hold for boys and girls.

Method

Participants and Procedure

The sample was part of a large longitudinal study on intrapersonal developmental risk factors in childhood and adolescence based at the University of Potsdam, Germany. The children were recruited from 33 public primary schools in the Federal State of Brandenburg, Germany. At T1, the sample consisted of $N = 1{,}652$ children (52.06% girls) aged between 6 and 11 years ($M = 8.36$, $SD = 0.93$)[2]. At T2, 1,611 children (51.8% girls) participated again ($M = 9.12$ years, $SD = 0.93$, range 7.11-11.89) and at T3, the remaining sample consisted of 1,501 children (51.5% girls; $M = 11.07$ years, $SD = 0.92$, range 9.12-13.76). This corresponds to a high retention rate of 97.5% from T1 to T2 and 92.3% from T2 to T3.
The mean interval between T1 and T2 was 9.14 months ($SD = 1.80$), and between T2 and T3, it was 23.83 months ($SD = 1.66$).

Approval for the procedure and the instruments was granted by the Ethics Committee of the authors' university as well as the Ministry of Education, Youth, and Sport of the Federal State of Brandenburg. The EF tasks were administered in individual test sessions by trained project members at the participants' schools. Parents and teachers completed the questionnaires either online or in paper-pencil form. For each child, informed consent was obtained from the parents.

[2] The original sample size was $N = 1{,}657$, however in the present study, 5 children were excluded due to missing values on the cluster variable that was included in the analyses (see section 2.3. Plan of Analyses)

Study 2

Measures

Executive function

The EF subcomponent *inhibition* was assessed by the Fruit Stroop task (Archibald & Kerns, 1999; adapted by Röthlisberger, Neuenschwander, Michel, & Roebers, 2010), a child-version of a Stroop paradigm with vegetables and fruits as stimulus items. The task consisted of four trials, and in each a page depicting 25 stimuli was presented to the child. Page 1 depicted colored rectangles (blue, green, red, yellow), page 2 showed fruits or vegetables in their typical colors (plum – blue, lettuce – green, strawberry – red, banana – yellow). Page 3 depicted the same fruits and vegetables, but all were colored grey. Page 4 displayed the same fruits and vegetables, but all were colored incorrectly. The child was instructed to name the correct color of the stimuli (pages 1 and 2), or to name the color that the fruits and vegetables should have (pages 3 and 4), as quickly as possible. For each page, the time (in seconds) required for giving correct responses for all 25 stimuli was measured. As dependent variable, an interference score was generated based on Röthlisberger et al. (2010): time p.4 - [(time p.1 × time p.3)/(time p.1 + time p.3)]. Higher scores indicated a lower ability to successfully inhibit the prepotent response of naming the color in which the stimuli were depicted on page 4.

The EF subcomponent *working memory* was assessed using the Digit-Span Backward task (Petermann & Petermann, 2007). This is a complex working memory task (Best & Miller, 2010), and measures of complex working memory and updating have been found to be highly correlated in children (e.g., St Clair-Thompson & Gathercole, 2006). In this task, participants were told a sequence of digits, which they had to repeat in reverse order. Each trial consisted of two sequences of equal length. The first two sequences were 2 digits long, and in each of the next trials, the sequences were lengthened by one digit up to a maximum number of eight digits, yielding a total of 7 trials with 14 sequences. Within each trial, at least one sequence had to be answered correctly in order to proceed to the next trial. The dependent variable was the total number of sequences that had been repeated correctly with a potential range of 0 to 14.

The EF subcomponent *shifting* was assessed using the Cognitive Attention Shifting task (Röthlisberger et al., 2010; adapted from Zimmermann, Gondan, & Fimm, 2002). Participants were presented with a single-colored fish and a multi-colored fish appearing simultaneously on the left- and right-hand side, respectively, of the computer screen. Children were told to feed each kind of fish and to always alternate between the two kinds by pressing one of two keys on a QWERTZ keyboard. Across several trials, the side on which the two kinds of fish appeared changed randomly. This required the children to remember their previous response – that is, which kind of fish they fed – in order to maintain the requirement of alternating feeding. A total of 46 trials (interstimulus intervals ranged from 300 to 700 ms) was separated by a short break during which positive feedback was given. The dependent variable was the number of correct responses for the 22 switch trials, that is, the trials that required children to change their response pattern (i.e., from alternately pressing left/right to repeating left/left or right/right; Austin, Groppe, & Elsner, 2014).

The EF subcomponent *planning* was measured using items of the Planning and Organizing-scale from the Behavior Rating Inventory of Executive Function (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000). Eight of the original 10 items were selected based on their factor loadings and translated into German by two native speakers. The items covered a range of problems that students can face when they need to plan or organize present and future tasks for school (e.g., “does not plan tasks for school in advance”). Teachers indicated planning disability of their students during the past 6 months using a 5-point response scale ranging from 1 (*never)* to 5 (*always)*. A total score was computed by averaging the item scores. The internal consistency was high with $\alpha = .93$.

Study 2

Aggression

At all three time points, aggression was measured using a teacher-report questionnaire that contained subscales with 3 items each for physical and relational aggression as well as for proactive and reactive aggression. The teachers first rated the frequency of physical and relational aggression during the past 6 months on a 5-point scale ranging from 1 (*never)* to 5 (*daily)* (physical aggression: e.g., "hit, shoved, or pushed peers"; relational aggression: e.g., "spread rumors or gossips about some peers"). The items were adapted from the Children's Social Behavior Scale – Teacher Form (CSBS-T; Crick, 1996). In a next step, teachers were asked to rate the *functions* of aggressive behaviors, based on the Instrument of Reactive and Proactive Aggression (IRPA; Polman, De Castro, Thomaes, & van Aken, 2009; proactive aggression: e.g., "to be the boss", reactive aggression: e.g., "because someone teased or upset him/her"). The response scale ranged from 1 (*never*) to 5 (*always*). The items on the function of aggression were only completed if the total score of the frequency of physical and relational aggression was larger than 1. Thus, the children for whom the teachers reported no physical or relational aggression at all had logical missing values on the measures of proactive and reactive aggression. The handling of these missing values is explained below.

For all four subscales, total scores were created by averaging the corresponding items, based on acceptable to high internal consistencies (physical aggression: α_{t1} = .93, α_{t2} = .94, α_{t3} = .93; relational aggression: α_{t1} = .91, α_{t2} = .92, α_{t3} = .91; proactive aggression: α_{t1} = .80, α_{t2} = .77, α_{t3} = .81; reactive aggression: α_{t1} = .85, α_{t2} = .84, α_{t3} = .88).

Habitual Anger

At T1 and T2, parents rated the extent to which their children habitually experienced anger with the subscale Anger/Frustration of the Temperament in Middle Childhood Questionnaire (TMCQ; Simonds, 2006). This subscale assesses the extent of anger shown by the child in response to the interruption of ongoing tasks or goal blocking (e.g., "my child gets angry when she or he has trouble with a task“). The scale contains 7 items, and the response scale ranges from 1 (*almost always untrue*) to 5 (*almost always true*). A total score was obtained by averaging the item scores. The internal consistency was good with α = .80 at both time points.

Information-Processing Capacity

Information-processing capacity was assessed at T1with the Digit-Symbol Test (DST) of the German version of the Wechsler Intelligence Scale for Children (Petermann & Petermann, 2007). Children were given a worksheet on which they had to assign either common shapes (Version A; ages 6-7) or the numbers 1 to 7 (Version B; ages 8 and older) to various symbols. A key in which a specific shape/number was paired with each of the symbols was presented in the first row of the worksheet. For both versions, the number of correct symbols allocated within 120s was measured (standardized T-values were calculated). Information-processing capacity was measured to control for basic intellectual ability, which could be confounded with EF.

Plan of Analysis

SPSS (Version 23) was used for descriptive computations, and the hypotheses were analyzed through structural equation models using Mplus (Version 7.4; Muthén & Muthén, 2012). In all models, the robust Maximum Likelihood estimator (MLR) was used to account for the non-normal distribution of the data. Missing data were handled by the Full Information Maximum Likelihood (FIML) estimation option to avoid a reduction in sample size. To be able to use the FIML approach for the logical missings on the items of the functions of aggression, we included a participant’s overall frequency scores of aggression at all three time points as covariates in the models. The frequency of aggression is a perfect predictor of the presence or absence of a data point on the two functions of aggression. Therefore, missing data could be treated as missing at random, which allowed us to use the FIML approach (Enders, 2010).

Study 2

Because participants were nested within school classes, class was included as a cluster variable in all analyses. Due to the trait-like nature of aggressive behavior, we included a random intercept for both forms and both functions of aggression, following the recommendation of Hamaker, Kuiper, and Grasman (2015). Because habitual anger and the three EF subcomponents were not assessed at all three time points, we were not able to include random intercepts for these variables (at least three time points are required to specify random intercepts). EF was modeled as a latent factor using the measures of *working memory*, *inhibition*, *shifting* and *planning* as indicators.

The model fits of the measurement model of EF and of the structural equation models were evaluated based on the criteria of Hu and Bentler (1998), with a comparative fit index (CFI) $\geq .95$, a root-mean-square error of approximation (RMSEA) $\leq .06$, and a standardized root-mean-square residual (SRMR) $\leq .08$ indicating a good fit. The χ^2-statistic was not interpreted as a measure of absolute fit, because it is biased in large samples (Schermelleh-Engel, Moosbrugger, & Müller, 2003). Bootstrap analyses were used to test indirect effects. If the bootstrapped 95% confidence interval does not include zero, the indirect effect is considered to be significant (Shrout & Bolger, 2002). The potential moderating effect of gender was examined using multi-group analyses. The measurement invariance between the gender groups was assessed based on comparisons between a fully constrained and a fully unconstrained (freed) model. The indicator for measurement invariance was a nonsignificant difference in χ^2 with scale corrections for the MLR estimator, as proposed by Satorra and Bentler (2001) or a nonsignificant Wald test for invariance in the indirect effects.

Results

Descriptive Statistics, Gender Differences, Factor Analysis, and Correlations

The descriptive statistics of all study variables are presented in Table 1. Gender differences were analyzed using *t*-tests for independent samples. If the assumption of homogeneity of variance was violated (as indicated by the Levene's Test for Equality of Variances), the degrees of freedom were adjusted using the Welch-Satterthwaite method. To account for multiple testing, we used a strict alpha level of $p < .01$. Effect size was calculated as Cohen's *d*. Boys were rated to be significantly more physically aggressive than were girls at all time points, all $ts \geq 11.00$, $ps < .001$, $ds \geq 0.60$. For relational aggression, no significant gender differences were found. Boys as compared to girls were also rated to show significantly more proactive aggression at T1 and T3, all $ts \geq 3.48$, $ps < .01$, $ds \geq 0.25$, and more reactive aggression at T2 and T3, all $ts \geq 3.48$, $ps < .01$, $ds \geq 0.30$. For the measure of anger, a significant gender difference was found only at T1, with higher scores for boys than for girls, $t(1332) = 3.64$, $p < .001$, $d = 0.19$.

With regard to EF, girls scored significantly higher than did boys on shifting, $t(1535) = 3.79$, $p < .001$, $d = 0.19$, and boys scored significantly higher than did girls on inhibition, $t(1534.04) = 3.24$, $p < .01$, $d = 0.16$, and planning, $t(1415) = 8.23$, $p < .001$, $d = 0.56$. Furthermore, girls scored significantly higher than did boys on information-processing capacity, $t(1642) = 7.18$, $p < .001$, $d = 0.35$. Due to these differences, gender was included as a covariate in the structural equation models (except for the multi-group model).

A latent factor of EF was specified by using the measures of working memory, inhibition, shifting, and planning as manifest indicators. Factor loadings were of moderate size (working memory: 0.55, inhibition: -0.62, shifting: 0.46, planning: -0.51; all $ps < .001$). The resulting measurement model showed a good fit ($\chi^2[3] = 2.76$, $p = .25$; CFI = 1.00; RMSEA = .02, 90% CI [.00, .05]; SRMR = .01).

Table 2 presents the correlations among all study variables and their links with age and information-processing capacity. The following significant correlations were found: EF at T1 was negatively correlated with anger at T1 and T2, and with all aggression measures at T1, T2, and T3. Anger was positively correlated with all aggression measures within and across time points. Age and information-processing capacity were positively correlated with EF.

Furthermore, age was negatively correlated with reactive aggression at T3, and information-processing capacity was negatively correlated with physical aggression at all time points and with reactive aggression at T1 and T3. As a consequence, we decided to include age and information-processing capacity as covariates in all models.

Hypothesis-Testing Analyses

Two separate models were specified to examine the proposed links between EF, anger, and aggression: one for the forms of aggression (physical and relational; see Figure 1) and one for the functions of aggression (reactive and proactive; see Figure 2). Age, gender, and information-processing capacity were included as covariates in both models.

Links between EF, Anger, and Forms of Aggression

The model for the forms of aggression (Figure 1) showed an acceptable model fit (χ^2[39] = 363.05, $p < .001$; CFI = .93; RMSEA = .07, 90% CI [.06, .08]; SRMR = .05). In line with Hypothesis 1, we found that controlling for stable individual differences in aggression, there was a significant negative path from EF at T1 to physical aggression at T3. Regarding relational aggression at T3, our data revealed a significant negative link to EF at T1 as well. Thus, the lower a child's EF, the higher the teacher-rated frequency of physical and relational aggression after the 3-year period. The paths from EF at T1 to physical and relational aggression at T2 were also negative and significant.

A significant negative link between EF at T1 and habitual anger at T2 was found, indicating that the lower children's EF was at T1, the more anger-prone they were rated by their parents at T2. Moreover, there was a significant positive link between habitual anger at T1 and both physical and relational aggression at T2, and there was a significant positive path from habitual anger at T2 to physical aggression at T3, but no significant link to relational aggression at T3.

Study 2

Hypothesis 2 postulated indirect negative effects between EF at T1 and both forms of aggression at T3 through habitual anger at T2. This hypothesis was only partially confirmed, because an indirect effect was found only for physical aggression, $\beta = -.01$, 95% CI [-.021, -.001]. For relational aggression, the indirect path was not found due to the nonsignificant path from habitual anger at T2 to relational aggression at T3 Links between EF, Anger, and Functions of Aggression. The model for the functions of aggression (Figure 2) also showed an acceptable fit ($\chi^2[61] = 407.04$, $p < .001$; CFI = .91; RMSEA = .06, 90% CI [.05, .06]; SRMR = .05). As predicted in Hypothesis 3, we found that controlling for stable individual differences in aggression, there was a significant negative path from EF at T1 to reactive aggression at T3, whereas the path from EF at T1 to proactive aggression at T3 was nonsignificant. Similarly, there was a significant negative link between EF at T1 and reactive, but not proactive aggression at T2. Finally, Hypothesis 4 postulated an indirect negative effect between EF at T1 and reactive aggression at T3 through habitual anger at T2. This hypothesis was not supported, because no significant link between T2 habitual anger and T3 reactive aggression was found.

Multi-Group Analyses of Potential Gender Differences

To examine potential gender differences in our first model, considering physical and relational aggression, we compared a fully unconstrained model, in which all paths were allowed to vary between girls and boys (fit: $\chi^2[72] = 396.54$, $p < .001$; CFI = .92; RMSEA = .07, 90% CI [.07, .08]; SRMR = .05), with a fully constrained model, in which all paths were constrained to be equal (fit: $\chi^2[(132] = 650.12$, $p < .001$; CFI = .88; RMSEA = .07, 90% CI [.06, .07]; SRMR = .12). The difference in χ^2 was significant, $\Delta\chi^2(60) = 262.66$, $p < .001$, indicating gender differences in specific parts of the model. Therefore, we computed a revised model (fit: $\chi^2[121] = 442.26$, $p < .001$; CFI = .92; RMSEA = .06, 90% CI [.05, .06]; SRMR = .07), in which we constrained all paths of the model to be equal between boys and girls, but with free estimation of those covariances and intercepts that were found to be different between boys and girls (e.g., the covariances between physical and relational aggression at all three time-points). The revised model had a significantly better fit than the fully constrained model, $\Delta\chi^2(11) = 150.69$, $p < .001$, and did not fit significantly worse than the fully unconstrained model, $\Delta\chi^2(49) = 66.0$, $p = .053$. In the revised model, there were no significant gender differences in the hypothesized paths (all $\Delta\chi^2$s ≤ 3.64, *ps* ≥ .19; all *W*[1] ≤ 0.13, *ps* ≥ .72).

For our second model, considering reactive and proactive aggression, we followed the same procedure. We also compared a fully unconstrained model, in which all paths were allowed to vary between girls and boys (fit: $\chi^2[116] = 445.94$, $p < .001$; CFI = .91; RMSEA = .06, 90% CI [.05, .06]; SRMR = .05), with a fully constrained model (fit: χ^2([198] = 624.94, $p < .001$; CFI = .88; RMSEA = .05, 90% CI [.05, .06]; SRMR = .08). The difference in χ^2 was significant, $\Delta\chi^2(82) = 188.41$, $p < .001$. Then, we computed a revised model (fit: $\chi^2[191] = 524.07$, $p < .001$; CFI = .91; RMSEA = .05, 90% CI [.04, .05]; SRMR = .06), in which we constrained all paths of the model to be equal between boys and girls, but with free estimation of some covariances and intercepts that were found to be different (for instance the covariances between the overall frequency scores of aggression at the three time-points). The revised model had a significantly better fit than the fully constrained model, $\Delta\chi^2(7) = 86.93$, $p < .001$, and did not fit significantly worse than the fully unconstrained model, $\Delta\chi^2(75) = 91.87$, $p = .090$. In the revised model, there were no significant gender differences in the hypothesized paths (all $\Delta\chi^2$s ≤ 2.79; *ps* ≥ .095; all *W*[1] ≤ 0.63, *ps* ≥ .43).

Discussion

The aim of the present study was to examine the longitudinal associations of EF (calculated as a latent factor of EF from behavioral measures of inhibition, set shifting, and working-memory updating, as well as teacher-reported planning), parent-reported habitual anger, and teacher-reported forms of aggression (i.e., physical and relational) and functions of aggression (i.e., proactive and reactive) in middle childhood. The hypotheses were examined in a large population-based sample in a three-wave design over a period of three years.

In line with Hypothesis 1, we found that EF was a significant negative predictor of physical aggression, and we also found a significant negative path between EF and relational aggression. This was true for the paths from T1 to T2 as well as from T1 to T3. Thus, the more deficits in EF children showed at T1 the higher was their teacher-rated frequency of both forms of aggression 1 and 2 years later. Our results held after controlling for information-processing capacity, gender, and age in the whole model. Further, we controlled for stable between-person differences by inclusion of a random intercept for forms and functions of aggression. Following the reasoning by Hamaker et al. (2015), this method allowed us to uncover causal relationships in within-persons processes. Therefore, our findings replicate longitudinal findings from other age groups (e.g., Hughes & Ensor, 2008; Hughes et al., 1998; Ogilvie et al., 2011; Olson et al., 2017; Schoemaker et al., 2013), and they extend previous cross-sectional research in middle childhood (e.g., McQuade et al., 2013), showing that lower EF at a mean age of 8 years predicted higher physical and relational aggression at about 9 and 11 years as within-person change. Furthermore, the negative path from EF to physical aggression is consistent with the frontal-lobe hypothesis of physical aggression (Séguin, 2009), and the social information-processing theory of aggression (Crick & Dodge, 1994). Consequently, the children in our study, who showed physically and relationally aggressive behavior at the age of 9 and 11 years might already have manifested significant cognitive deficits in their EF abilities, located in the frontal lobe, at the age of 8 years.

Our finding of significant negative paths from EF at T1 to relational aggression at both T2 and T3 confirms some of the previous cross-sectional findings (McQuade et al., 2013), but contradicts others (no significant path: Diamantopoulou et al., 2007; positive relation: Poland et al., 2015). Nevertheless, our finding is in line with the social information-processing theory of aggression that proposes that physically as well as relationally aggressive children have deficits in their cognitive processing of social situations (Crick & Dodge, 1994). To further examine the role of EF in the development of relational aggression, it may be important to include a more differentiated assessment of relational aggression that considers the complexity of different relationally aggressive behaviors. According to Crick et al. (1999), relationally aggressive behaviors can take different forms ranging from relatively simple, direct types (e.g., threatening to end the friendship) to more complex, indirect types (e.g., mobilizing peer group members against a certain child to make that child feel excluded). The latter requires a higher level of cognitive skills to be used effectively. Thus, it may be that only the direct types of relational aggression are related to deficits in EF, whereas the indirect types of relational aggression are unrelated or even positively related to EF. However, this remains to be tested in future studies.

In Hypothesis 2, we postulated that anger would mediate the link between EF and both forms of aggression. However, this prediction was only confirmed for physical, and not for relational aggression. For physical aggression, the results indicate that the lower the children's EF at T1 the higher their parent-reported habitual anger was at T2, which in turn predicted more teacher-rated physical aggression at T3. Even though the size of the indirect effect was small, it was still significant after controlling for age, gender, information-processing capacity, and stable between-person differences of physical and relational aggression. The tendency to experience anger is only one of many intrapersonal factors involved in the complex emergence of aggressive behavior (see Krahé, 2013, for a review). However, our findings extend previous cross-sectional research (e.g., De Castro et al., 2005) and support the role of anger as one explanatory construct in the link between EF and physical aggression in middle childhood. Further, the negative mediation between EF and physical aggression by anger highlights the importance of considering emotions in the social-cognitive information processing of children who display aggressive behavior (Lemerise & Arsenio, 2000). Our study is – as far as we know – the first to uncover this meditational effect over a time-span of three years in middle childhood, which underlines the need to consider large time intervals in the development of physical aggression, as well as both emotional and cognitive processes.

For relational aggression, this mediation effect was not found. This finding is inconsistent with theory and previous research that has found anger to be involved in the development of both physical and relational aggression (Crick et al., 1999; Crick, Grotpeter, & Bigbee, 2002). However, the positive path from anger at T2 to relational aggression at T3 only narrowly missed the level of significance ($p = .054$, $\beta = .06$). This trend tentatively suggests that the tendency to experience anger may be involved in the negative link between EF and later relational aggression. To date, only few studies have differentiated between forms of aggression when examining the link between anger and aggression. Future research is needed to explore potential differences between relational and physical aggression regarding the association with EF and anger, not only in middle childhood.

With regard to the functions of aggression, we found that EF was a significant negative predictor of reactive aggression, but that EF was unrelated to proactive aggression. This pattern is consistent with Hypothesis 3 and was found for the paths from EF at T1 to functions of aggression both at T2 and at T3. It is in line with the few previous cross-sectional studies regarding the association between EF and the functions of aggression in middle childhood (e.g., Ellis et al., 2009; White et al., 2013). Moreover, the nonsignificant path from EF to proactive aggression is consistent with research on CU traits, which are conceptually linked to proactive aggression (e.g., Pardini et al., 2003). Several studies have shown that CU traits are unrelated to deficits in different domains of EF such as inhibition (Tye et al. 2017) or set shifting (Mitchell, Colledge, Leonard, & Blair, 2002). Furthermore, our findings support the theoretical differentiation of the two functions of aggression. Higher EF enables children to behave in a planned and deliberate fashion, which is characteristic of proactive aggression. In contrast, reactive aggression refers to impulsive aggressive acts that do not require sophisticated planning. Thus, the inability to plan and to inhibit behavioral responses, both components of low EF, may explain the direct negative paths between EF and reactive aggression at one or two years later found in this study.

Contrary to Hypothesis 4, no indirect link between EF and reactive aggression via anger was found. We did find a negative link between EF and anger, confirming previous evidence that EF is involved in the regulation of negative affect in children (e.g., Carlson & Wang, 2007). This finding is also consistent with research on deficits in processes associated with EF, for instance error monitoring, in children with chronic irritability (Perlman et al., 2015). However, there was no link between anger and reactive aggression, which is surprising given that anger is assumed to be a major impelling factor of reactive aggression during childhood (e.g., Arsenio et al., 2000; Eisenberg et al., 1999; Leibenluft & Stoddard, 2013; Olson et al., 2005). One explanation may be that the relation between anger and reactive aggression depends on moderating variables. These may include the ability to regulate the behavioral impulses related to anger, or the tendency to act impulsively, which relates to less voluntary and more reactive aspects of control (Rothbart & Rueda, 2005). Another reason could be that until now the relations of chronic anger, irritability, and disruptive behavior symptoms were mainly investigated in clinical samples with no differentiation between reactive and proactive aggression (e.g., Wakschlag et al., 2015). In addition, further variables besides anger may act as mediators of the link between EF and reactive aggression. These may include social information-processing variables, such as theory of mind (Renouf, Brendgen, Séguin, et al., 2010) or hostile attribution bias (De Castro, Veerman, Koops, Bosch, & Monshouwer, 2002).

In addition to the examination of the longitudinal links between EF and aggression, our study contributes to previous research by considering potential gender differences. We did find gender differences on some of the study variables, in particular on the frequency of physical aggression, with boys scoring higher than girls at all three time points. This finding is in line with a meta-analysis on gender differences in aggressive behavior (Card et al., 2008). However, the multi-group analyses revealed that the predictive paths from EF to aggression and from EF to habitual anger as well as the indirect paths from EF over anger to aggression did not vary by gender, which confirms and extends previous cross-sectional research (White et al., 2013). Consequently, the processes and mechanisms that lead from EF to aggressive behavior seem to be equivalent in girls and boys in middle childhood.

Strengths and Limitations

We believe that the present study has several strengths. These include the large sample size, the longitudinal design with three time points covering about three years across middle childhood, the differentiation of forms and functions of aggression, and the examination of potential gender differences. Furthermore, we used three different sources to assess the study constructs, a procedure that is known to reduce common method biases (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). A final strength is the inclusion of a random intercept, by which we controlled for the stability of individual differences in the forms and functions of aggressive behavior. This new procedure is recommended to overcome the limitations of the traditional cross-lagged model by separating within-person change, which is the focus of our models, from stable between-person differences (Hamaker et al., 2015).

Despite these strengths, some limitations have to be noted. One refers to the distinction between “cool” and “hot” aspects of EF during childhood and adolescence (Zelazo & Carlson, 2012). Cool EF is usually associated with the lateral prefrontal cortex and operates in relation to more abstract and decontextualized problems. In contrast, hot EF is associated with the orbitofrontal cortex and operates in more motivationally and emotionally significant situations (Zelazo & Müller, 2002). The measures we used to assess EF covered only the cool component. Nevertheless, our findings demonstrate that cool EF alone seems to play an important role in the prediction of physical, relational, and reactive aggression in middle childhood. A further limitation of our study is that aggression was only assessed by teacher reports. This may have led to an underestimation of relationally aggressive behavior because it includes more indirect forms of aggression that may be less obvious for teachers. Thus, peer reports or peer nominations could provide important additional information because aggressive behavior usually takes place within the peer group. Finally, it is important to mention that our sample was a community sample. Whether similar associations are also found for clinical populations of youth with serious levels of aggression and/or chronic symptoms of irritability is a question for future research.

Conclusion and Implications

Our study extends the existing literature about the relation between EF and aggression in middle childhood by taking the habitual tendency to experience anger as a potential mediator into account, by considering different forms and functions of aggressive behavior, and by analyzing gender differences. We found that EF predicted physical, relational, and reactive aggression over the course of middle childhood, and that the link between EF and physical aggression was partly mediated by habitual proneness to anger. Although there were gender differences in the frequency of aggressive behavior, these predictive paths were not moderated by gender. Our findings highlight the importance of addressing EF in programs that aim to reduce aggressive behavior. In the last years, an increasing number of programs to promote EF has been developed, and the effectiveness of these programs has been demonstrated (Diamond & Lee, 2011), making EF a promising candidate for the prevention of aggression. For the prevention of physical aggression in particular, teaching strategies for coping with anger should also be considered. Regarding gender, our findings indicate that both gender groups should be addressed in prevention programs. Although the mean level of physically aggressive behavior in particular may be higher among boys, the paths between EF, anger, and aggression seem to be similar for girls and boys. Our study underlines the need to promote the development of EF not only in early, but also in middle childhood to prevent later physical, relational, and reactive aggression and its negative consequences.

Author Contributions

All authors have contributed substantially to the conceptualization and the design of the work. HR, AH, and FK: Have primarily collected, analyzed and interpreted the data; HR and AH: Have written the first draft of the paper, and all authors contributed to revise the paper; BK and BE: Have provided input and supervision to the analyses. All authors agreed to be accountable for all aspects of the work.

Funding

This research was funded by the German Research Foundation as part of the Research Training Unit “Intrapersonal developmental risk factors in childhood and adolescence: A longitudinal perspective” (GRK 1668). We also acknowledge financial support from the Open Access Publishing Fund of the University of Potsdam.

Study 2

References

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, *8*, 71–82. http://doi.org/10.1076/chin.8.2.71.8724

Archibald, S. J., & Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychology*, *5*, 115–129. http://doi.org/10.1076/chin.5.2.115.3167

Arsenio, W. F., Cooperman, S., & Lover, A. (2000). Affective predictors of preschoolers' aggression and peer acceptance: Direct and indirect effects. *Developmental Psychology*, *36*, 438–448. http://doi.org/10.1037//0012-1649.36.4.438

Austin, G., Groppe, K., & Elsner, B. (2014). The reciprocal relationship between executive function and theory of mind in middle childhood: a 1-year longitudinal perspective. *Frontiers in Psychology*, *5*(655), 1–11. http://doi.org/10.3389/fpsyg.2014.00655

Baron, R. A., & Richardson, D. R. (1994). *Human aggression*. New York: Plenum Press.

Berkowitz, L., & Harmon-Jones, E. (2004). Toward an understanding of the determinants of anger. *Emotion*, *4*, 107–130. http://doi.org/10.1037/1528-3542.4.2.107

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, *81*, 1641–1660. http://doi.org/10.1111/j.1467-8624.2010.01499.x

Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review*, *29*, 180–200. http://doi.org/10.1016/j.dr.2009.05.002

Blair, R., Leibenluft, E. and Pine, D. (2014). Conduct disorder and callous–unemotional traits in youth. *New England Journal of Medicine*, *371*, 2207-2216. http://doi.org/101056/NEJMral1315612

Blakemore, S.-J., & Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry,* *and Allied Disciplines*, *47*, 296–312. http://doi.org/10.1111/j.1469-7610.2006.01611.x

Bridgett, D. J., Oddi, K. B., Laake, L. M., Murdock, K. W., & Bachmann, M. N. (2013). Integrating and differentiating aspects of self-regulation: effortful control, executive functioning, and links to negative affectivity. *Emotion*, *13*, 47–63. http://doi.org/10.1037/a0029536

Card, N. A., & Little, T. D. (2006). Proactive and reactive aggression in childhood and adolescence: A meta-analysis of differential relations with psychosocial adjustment. *International Journal of Behavioral Development*, *30*, 466–480. http://doi.org/10.1177/0165025406071904

Card, N. A., Stucky, B. D., Sawalani, G. M., & Little, T. D. (2008). Direct and indirect aggression during childhood and adolescence: a meta-analytic review of gender differences, intercorrelations, and relations to maladjustment. *Child Development*, *79*, 1185–1229. http://doi.org/10.1111/j.1467-8624.2008.01184.x

Carlson, S. M., & Wang, T. S. (2007). Inhibitory control and emotion regulation in preschool children. *Cognitive Development*, *22*, 489–510. http://doi.org/10.1016/j.cogdev.2007.08.002

Côté, S. M., Vaillancourt, T., Barker, E. D., Nagin, D., & Tremblay, R. E. (2007). The joint development of physical and indirect aggression: Predictors of continuity and change during childhood. *Development and Psychopathology*, *19*, 37–55. http://doi.org/10.1017/S0954579407070034

Crick, N. R. (1996). The role of overt aggression, relational aggression, and prosocial behavior in the prediction of children’s future social adjustment. *Child Development*, *67*, 2317–2327. http://doi.org/10.1111/j.1467-8624.1996.tb01859

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin*, *115*, 74–101. http://doi.org/10.1037/0033-2909.115.1.74

Crick, N. R., & Grotpeter, J. K. (1995). Relational aggression, gender, and social-psychological adjustment. *Child Development*, *66*, 710–722. http://doi.org/10.2307/1131945

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*, 1134–1142. http://doi.org/0009-3920/2002/7304-0010

Crick, N. R., Werner, N. E., Casas, J. F., O´Brien, K. M., Nelson, D. A., Grotpeter, J. K., & Markon, K. (1999). Childhood aggression and gender: A new look at an old problem. In D. Bernstein (Ed.), *Nebraska symposium on motivation* (pp. 75–141). Lincoln, NE: University of Nebraska Press.

Cui, L., Colasante, T., Malti, T., Ribeaud, D., & Eisner, M. P. (2016). Dual trajectories of reactive and proactive aggression from mid-childhood to early adolescence: Relations to sensation seeking, risk taking, and moral reasoning. *Journal of Abnormal Child Psychology*, *44*, 663–675. http://doi.org/10.1007/s10802-015-0079-7

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*, 105–116. http://doi.org/10.1207/s15374424jccp3401

de Castro, B. O., & van Dijk, A. (2018). "It's gonna end up with a fight anyway". Social cognitive processes in children with disruptive behavior disorders. In J. E. Lochman & W. Matthys (Eds.), *The Wiley Handbook of disruptive and impulse-control disorders* (pp. 237–253). Hoboken, NJ: Wiley.

Study 2

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*, 916–934. http://doi.org/10.1111/1467-8624.00447

Diamantopoulou, S., Rydell, A. M., Thorell, L. B., & Bohlin, G. (2007). Impact of executive functioning and symptoms of attention deficit hyperactivity disorder on children's peer relations and school performance. *Developmental Neuropsychology*, *32*, 521–542. http://doi.org/10.1080/87565640701360981

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4-12 years old. *Science*, *333*(6045), 959–964. http://doi.org/10.1126/science.1204529

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

Eisenberg, N., Fabes, R. A., Murphy, B. C., Shepard, S., Guthrie, I. K., Mazsk, P., … Jones, S. (1999). Prediction of elementary school children's socially appropriate and problem behavior from anger reactions at age 4-6 years. *Journal of Applied Developmental Psychology*, *20*, 119–142. http://doi.org/10.1016/S0193-3973(99)80007-0

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*, 945–956. http://doi.org/10.1007/s11103-011-9767-z.Plastid

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Press.

Gagne, J. R., & Hill Goldsmith, H. (2011). A longitudinal analysis of anger and inhibitory control in twins from 12 to 36 months of age. *Developmental Science*, *14*(1), 112–124. http://doi.org/10.1111/j.1467-7687.2010.00969.x

Study 2

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior Rating of Executive Function.* Lutz, FL: Psychological Assessment Resources.

Granvald, V., & Marciszko, C. (2016). Relations between key executive functions and aggression in childhood. *Child Neuropsychology*, *22*(5), 537–555. http://doi.org/10.1080/09297049.2015.1018152

Hamaker, E., Kuiper, R., & Grasman, R. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*, 102–116. http://dx.doi.org/10.1037/a0038889

Hawes, S. W., Perlman, S. B., Byrd, A. L., Raine, A., Loeber, R., & Pardini, D. A. (2016). Chronic anger as a precursor to adult antisocial personality features : The moderating influence of cognitive control. *Journal of Abnormal Child Psychology*, *125*, 64–74. http://doi.org/10.1037/abn0000129.Chronic

Healey, D. M., Marks, D. J., & Halperin, J. M. (2011). Examining the interplay among negative emotionality, cognitive functioning, and attention deficit/hyperactivity disorder symptom severity. *Journal of the International Neuropsychological Society*, *17*, 502–510. http://doi.org/10.1017/S1355617711000294

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424–453. http://doi.org/10.1037/1082-989X.3.4.424

Hubbard, J. A., Romano, L. J., McAuliffe, M. D., & Morrow, M. T. (2010). Anger and the reactive-proactive aggression distinction in childhood and adolescence. In M. Potegal, G. Stemmler, & C. D. Spielberger (Eds.), *International handbook of anger: constituent and concomitant biological, psychological, and social processes* (pp. 231–239). New York: Springer.

Huesmann, L. R., Eron, L. D., & Yarmel, P. W. (1987). Intellectual functioning and aggression. *Journal of Personality and Social Psychology*, *52*(1), 232–240. http://doi.org/10.1037/0022-3514.52.1.232

Hughes, C., Dunn, J., & White, A. (1998). Trick or treat?: uneven understanding of mind and emotion and executive dysfunction in “hard-to-manage” preschoolers. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *39*(7), 981–994. http://doi.org/9804031

Hughes, C., & Ensor, R. (2008). Does executive function matter for preschoolers’ problem behaviors? *Journal of Abnormal Child Psychology*, *36*, 1–14. http://doi.org/10.1007/s10802-007-9107-6

Hughes, C., Graham, A., & Grayson, A. (2005). Executive function in childhood: Development and disorder. In I. J. Oates (Ed.), *Cognitive development* (pp. 205–230). Milton Keynes, UK: Open University Press.

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*(11), 2017–2036. http://doi.org/10.1016/j.neuropsychologia.2006.01.010

Krahé, B. (2013). *The social psychology of aggression* (2nd ed.). New York: Psychology Press.

Lehto, J. E., Juujarvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*, 59–80. http://doi.org/10.1348/026151003321164627

Leibenluft, E. (2011). Severe mood dysregulation, irritability, and the diagnostic boundaries of bipolar disorder in youths. *American Journal of Psychiatry*, *168*, 129–142. http://doi.org/10.1176/appi.ajp.2010.10050766

Leibenluft, E. (2017). Irritability in children: what we know and what we need to learn. *World*

Psychiatry, *16*(1), 100–101. http://doi.org/10.1002/wps.20397

Leibenluft, E., & Stoddard, J. (2013). The developmental psychopathology of irritability. *Development and Psychopathology*, *25*, 1473–1487. http://doi.org/10.1017/S0954579413000722

Lemerise, E. A., & Arsenio, W. F. (2000). An integrated model of emotion processes and cognition in social information processing. *Child Development*, *71*, 107–118. http://doi.org/10.1111/1467-8624.00124

McQuade, J. D., Murray-Close, D., Shoulberg, E. K., & Hoza, B. (2013). Working memory and social functioning in children. *Journal of Experimental Child Psychology*, *115*(3), 422–435. http://doi.org/10.1016/j.jecp.2013.03.002

Mitchell, D. G. V., Colledge, E., Leonard, A., & Blair, R. J. R. (2002). Risky decisions and response reversal: Is there evidence of orbitofrontal cortex dysfunction in psychopathic individuals? *Neuropsychologia, 40*, 2013–2022.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100. http://doi.org/10.1006/cogp.1999.0734

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user’s guide. Statistical analysis with latent variables* (7th ed.). Los Angeles, CA: Muthén & Muthén.

O’Toole, S. E., Monks, C. P., & Tsermentseli, S. (2017). Executive function and theory of mind as predictors of aggressive and prosocial behavior and peer acceptance in early childhood. *Social Development*, *26*, 907–920. http://doi.org/10.1111/sode.12231

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

Olson, S. L., Choe, D. E., & Sameroff, A. J. (2017). Trajectories of child externalizing problems between ages 3 and 10 years: Contributions of children's early effortful control, theory of mind, and parenting experiences. *Development and Psychopathology*, *29*, 1333–1351. http://doi.org/10.1017/S095457941700030X

Olson, S. L., Sameroff, A. J., Kerr, D. C. R., Lopez, N. L., & Wellman, H. M. (2005). Developmental foundations of externalizing problems in young children: The role of effortful control. *Development and Psychopathology*, *17*, 25–45. http://doi.org/10.1017/s0954579405050029

Pardini, D.A., Lochman, J.E., & Frick, P.J. (2003). Callous/unemotional traits and social-cognitive processes in adjudicated youths. *Journal of the American Academy of Child and Adolescent Psychiatry, 42,* 364–371. http://doi.org/10.1097/00004583-200303000-00018

Perlman, S. B., Jones, B. M., Wakschlag, L. S., Axelson, D., Birmaher, B., & Phillips, M. L. (2015). Neural substrates of child irritability in typically developing and psychiatric populations. *Developmental Cognitive Neuroscience*, *14*, 71–80. http://doi.org/10.1016/j.dcn.2015.07.003

Petermann, F., & Petermann, U. (Eds.) (2007). *Hamburg-Wechsler-Intelligenztest für Kinder - IV (HAWIK-IV).* Bern, Switzerland: Huber.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *The Journal of Applied Psychology*, *88*, 879–903. http://doi.org/10.1037/0021-9010.88.5.879

Poland, S. E., Monks, C. P., & Tsermentseli, S. (2016). Cool and hot executive function as predictors of aggression in early childhood: Differentiating between the function and form of aggression. *British Journal of Developmental Psychology*, *34*(2), 181–197. http://doi.org/10.1111/bjdp.12122

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,* 183–193. http://doi.org/10.1007/s10802-008-9266-0

Putnam, S. P., Rothbart, M. K., & Gartstein, M. A. (2008). Homotypic and heterotypic continuity of fine-grained temperament during infancy, toddlerhood, and early childhood. *Infant and Child Development*, *17*, 387–405. http://doi.org/10.1002/icd.582

Rathert, J., Fite, P. J., & Gaertner, A. E. (2011). Associations between effortful control, psychological control and proactive and reactive aggression. *Child Psychiatry and Human Development*, *42*, 609–621. http://doi.org/10.1007/s10578-011-0236-3

Renouf, A., Brendgen, M., Séguin, J. R., Vitaro, F., Boivin, M., Dionne, G., … Pérusse, D. (2010). Interactive links between theory of mind, peer victimization, and reactive and proactive aggression. *Journal of Abnormal Child Psychology*, *38*(8), 1109–1123. http://doi.org/10.1007/s10802-010-9432-z

Rothbart, M. K., & Rueda, M. R. (2005). The development of effortful control. In U. Mayr, E. Awh, & S. Keele (Eds.), *Developing individuality in the human brain: A tribute to Michael I. Posner* (pp. 167-188). Washington, D.C.: American Psychological Association.

Röthlisberger, M., Neuenschwander, R., Michel, E., & Roebers, C. M. (2010). Exekutive Funktionen: Zugrundeliegende kognitive Prozesse und deren Korrelate bei Kindern im späten Vorschulalter [Executive functions: Underlying cognitive processes and their correlates in late preschool children]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *42*, 99–110. http://doi.org/10.1026/0049-8637/a000010

Study 2

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514. http://doi.org/10.1007/BF02296192

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*, 23–74. Retrieved from https://www.dgps.de/fachgruppen/methoden/mpr-online/

Schoemaker, K., Mulder, H., Deković, M., & Matthys, W. (2013). Executive functions in preschool children with externalizing behavior problems: A meta-analysis. *Journal of Abnormal Child Psychology*, *41*, 457–471. http://doi.org/10.1007/s10802-012-9684-x

Séguin, J. R. (2009). The frontal lobe and aggression. *European Journal of Developmental Psychology*, *6*, 100–119. http://doi.org/10.1080/17405620701669871

Séguin, J. R., Arseneault, L., & Tremblay, R. E. (2007). The development of self-regulation: Toward the integration of cognition and emotion. *Cognitive Development*, *22*(4), 401–405. http://doi.org/10.1016/j.cogdev.2007.08.008

Séguin, J. R., & Zelazo, P. D. (2005). Executive Function in early physical aggression. In R. E. Tremblay, W. W. Hartup, & J. Archer (Eds.), *Developmental origins of aggression* (pp. 307–329). New York: Guilford Press.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*, 422–445. http://doi.org/10.1037/1082-989X.7.4.422

Simonds, J. (2006). *The role of reward sensitivity and response: Execution in childhood extraversion*. Unpublished doctoral dissertation. University of Oregon.

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of*

Experimental Psychology, *59*, 745–759. http://doi.org/10.1080/17470210500162854

Tye, C., Bedford, R., Asherson, P., Ashwood, K. L., Azadi, B., Bolton, P., & McLoughlin, G. (2017). Callous-unemotional traits moderate executive function in children with ASD and ADHD: A pilot event-related potential study. *Developmental Cognitive Neuroscience*, *26*, 84–90. http://doi.org/10.1016/j.dcn.2017.06.002

Vitaro, F., Brendgen, M., & Barker, E. D. (2006). Subtypes of aggressive behaviors: A developmental perspective. *International Journal of Behavioral Development*, *30*(1), 12–19. http://doi.org/10.1177/0165025406059968

Wakschlag, L. S., Estabrook, R., Petitclerc, A., Henry, D., Burns, J. L., Perlman, S. B., … Briggs-Gowan, M. L. (2015). Clinical implications of a dimensional approach: The normal:abnormal spectrum of early irritability. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*(8), 626–634. http://doi.org/https://doi.org/10.1016/j.jaac.2015.05.016

Walters, G. D. (2005). Proactive and reactive aggression: A lifestyle view. In J. P. Morgan (Ed.), *Psychology of aggression* (pp. 29–43). Hauppauge, NY: Nova Science Publishers.

White, B. A., Jarrett, M. A., & Ollendick, T. H. (2013). Self-regulation deficits explain the link between reactive aggression and internalizing and externalizing behavior problems in children. *Journal of Psychopathology and Behavioral Assessment*, *35*, 1–9. http://doi.org/10.1007/s10862-012-9310-9

Zadeh, Y. Z., Im-Bolter, N., & Cohen, N. J. (2007). Social cognition and externalizing psychopathology: An investigation of the mediating role of language. *Journal of Abnormal Child Psychology*, *35*(2), 141–152. http://doi.org/10.1007/s10802-006-9052-9

Zelazo, P. D., & Carlson, S. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, *6*, 354-360.

Study 2

Zelazo, P. D., & Müller, U. (2002). Executive function in typical and atypical development. In U. Goswami (Ed.), *Handbook of childhood cognitive development* (pp. 445-469). Oxford: Blackwell.

Zimmermann, P., Gondan, M., & Fimm, B. (2002). KITAP Testbatterie zur Aufmerksamkeitsprüfung für Kinder [KITAP Attention Test Battery for Children]. Herzogenrath, Germany: Psychologische Testsysteme.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Study 2 – Table 1

Descriptive Statistics of the Study Variables for the Total Sample and for Boys and Girls

	n	Range	Total *M* (*SD*)	Boys *M* (*SD*)	Girls *M* (*SD*)
EF T1					
Updating	1635	0-16	6.18 (1.46)	6.21 (1.47)	6.16 (1.45)
Shifting	1555	0-22	18.16 (3.92)	**17.77 (3.94)**	**18.52 (3.86)**
Inhibition	1640	7-89[a]	24.91 8.76)	**25.65 (9.48)**	**24.24 (7.99)**
Planning	1417	1-5	2.30 (0.89)	**2.50 (0.90)**	**2.11 (0.85)**
Anger					
T1	1334	1-5	2.63 (0.74)	**2.71 (0.74)**	**2.56 (0.72)**
T2	1190	1-5	2.75 (0.71)	2.62 (0.72)	2.53 (0.70)
Forms of aggression					
T1 Physical A	1408	1-5	1.49 (0.79)	**1.79 (0.91)**	**1.21 (0.52)**
T2 Physical A	1145	1-5	1.48 (0.77)	**1.79 (0.90)**	**1.18 (0.46)**
T3 Physical A	1104	1-5	1.43 (0.75)	**1.67 (0.88)**	**1.20 (0.49)**
T1 Relational A	1405	1-5	1.50 (0.70)	1.54 (0.70)	1.47 (0.70)
T2 Relational A	1144	1-5	1.56 (0.76)	1.61 (0.79)	1.51 (0.73)
T3 Relational A	1102	1-5	1.50 (0.71)	1.55 (0.76)	1.45 (0.66)
Functions of aggression					
T1 Reactive A	756	1-5	2.11 (0.92)	2.81 (0.95)	2.63 (0.97)
T2 Reactive A	627	1-5	2.17 (0.89)	**3.00 (0.97)**	**2.63 (0.93)**
T3 Reactive A	568	1-5	2.00 (0.95)	**2.56 (1.09)**	**2.25 (1.00)**
T1 Proactive A	753	1-5	2.73 (0.96)	**2.21 (0.96)**	**1.98 (0.83)**
T2 Proactive A	621	1-5	2.85 0.98)	2.22 (0.90)	2.08 (0.86)
T3 Proactive A	568	1-5	2.42 (1.06)	**2.18 (0.99)**	**1.77 (0.81)**
DST T1	1644	27-80[a]	51.28 (9.17)	**49.61 (8.97)**	**52.81 (9.08)**

Notes: EF = executive function, A = aggression, DST = digit-symbol test; [a]Max values are theoretically infinite, thus, table values are sample-specific. Means in bold differ significantly between boys and girls.

Study 2

Table 2

Correlations between Executive Function, Anger, Aggression, and Age

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.
1. T1 EF		-.19***	-.23***	-.34***	-.33***	-.28***	-.24***	-.19***	-.24***	-.25***	-.27***	-.33***	-.15**	-.18**	-.16**	.42***	.45***
2. T1 Anger			.70***	.19***	.22***	.12***	.16***	.15***	.09***	.11**	.18**	.11**	.09*	.08	.08*	-.04	-.01
3. T2 Anger				.17***	.20***	.16***	.18***	.13***	.13***	.12**	.12***	.11*	.11**	.08*	.08*	-.02	-.06
4. T1 Phy A					.71***	.48***	.62***	.45***	.35***	.28***	.23***	.21***	.44***	.30***	.26***	-.13***	-.02
5. T2 Phy A						.50***	.44***	.65***	.37***	.19***	.28***	.26***	.34***	.40***	.28***	-.10***	-.03
6. T3 Phy A							.33***	.30***	.67***	.14***	.16***	.13***	.24***	.23***	.32***	-.12***	-.01
7. T1 Rel A								.53***	.37***	.19***	.13***	.13***	.40***	.25***	.19***	-.05	-.05
8. T2 Rel A									.30***	.11***	.21***	.17***	.30***	.39***	.21***	.02	-.04
9. T3 Rel A										.07**	.11***	.18***	.23***	.23***	.36***	-.04	-.02
10. T1 Reac A											.43***	.21***	.26***	.12***	.10*	-.06*	-.05
11. T2 Reac A												.36***	.18***	.19***	.15**	-.07	.04
12. T3 Reac A													.17***	.19***	.35***	-.11**	-.16**
13. T1 Proac A														.44*	.33*	.01	.02
14. T2 Proac A															.35*	-.01	.04
15. T3 Proac A																.02	-.00
16. DST																	.00
17. T1 Age																	

Notes: EF = executive function, Phy A = physical aggression, Rel A = relational aggression, Reac A = reactive aggression, Proac A = proactive aggression, DST

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

Study 2

Figure 1. Prediction of the Forms of Aggression

Notes: Standardized path coefficients are displayed; dotted lines = nonsignificant path coefficients; mean physical and mean relational aggression partial out between-person stability over time (random intercept); model fit: $\chi^2(39) = 363.05$, $p < .001$; CFI = .93; RMSEA = .07, 90% CI = [.06, .08]; SRMR = .05; for clarity of presentation, only paths between the time points are presented in the figure, but within-time correlations at all time points were also included in the model.

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

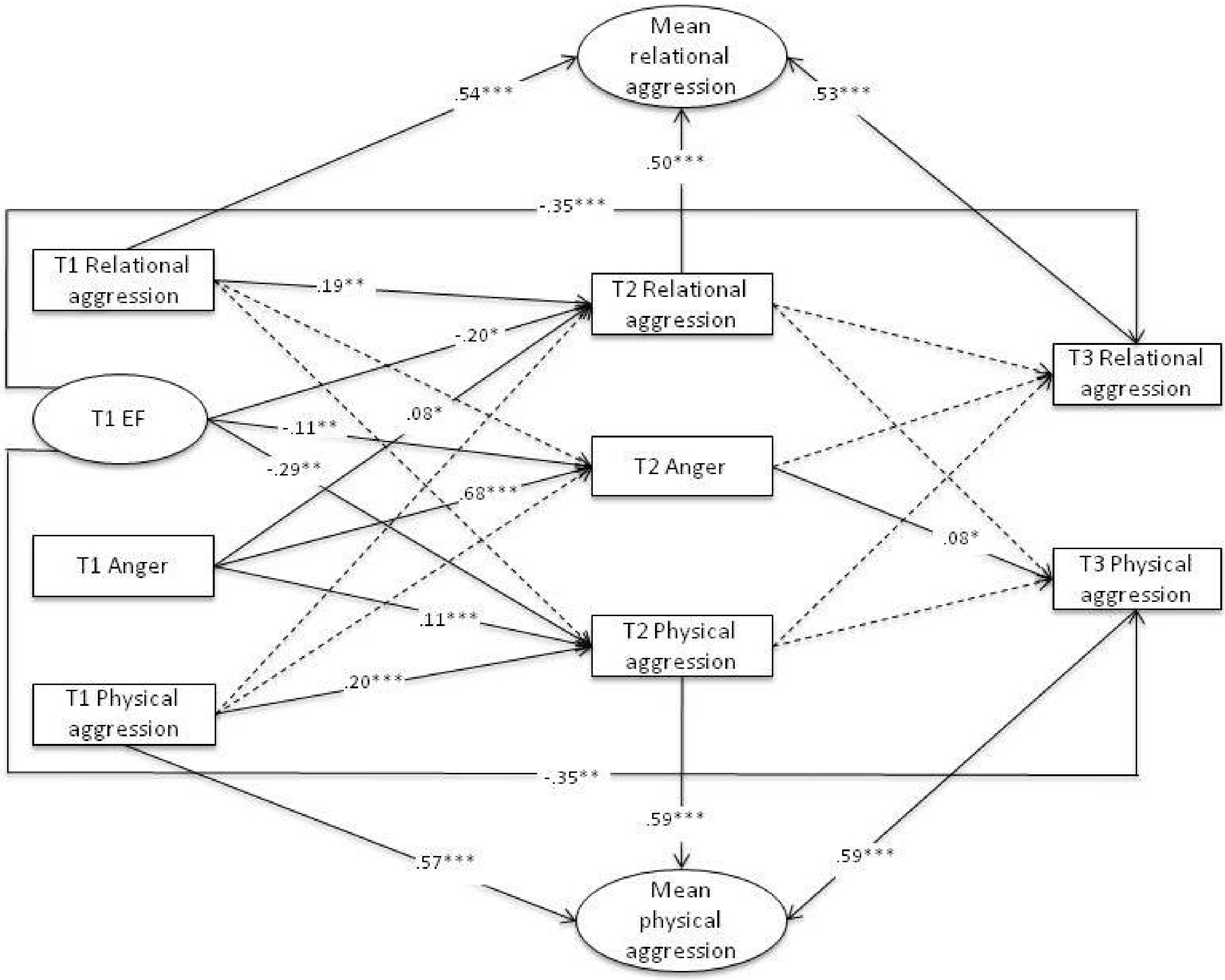
Figure 2. Prediction of the Functions of Aggression

Notes: Standardized path coefficients are displayed; dotted lines = nonsignificant path coefficients; mean proactive and mean reactive aggression partial out between-person stability over time (random intercept); model fit: $\chi^2(61) = 407.04$, $p < .001$; CFI = .91; RMSEA = .06, 90% CI = [.05, .06]; SRMR = .05; for clarity of presentation, only paths between the time points are presented in the figure, but within-time correlations at all time points were also included in the model.

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

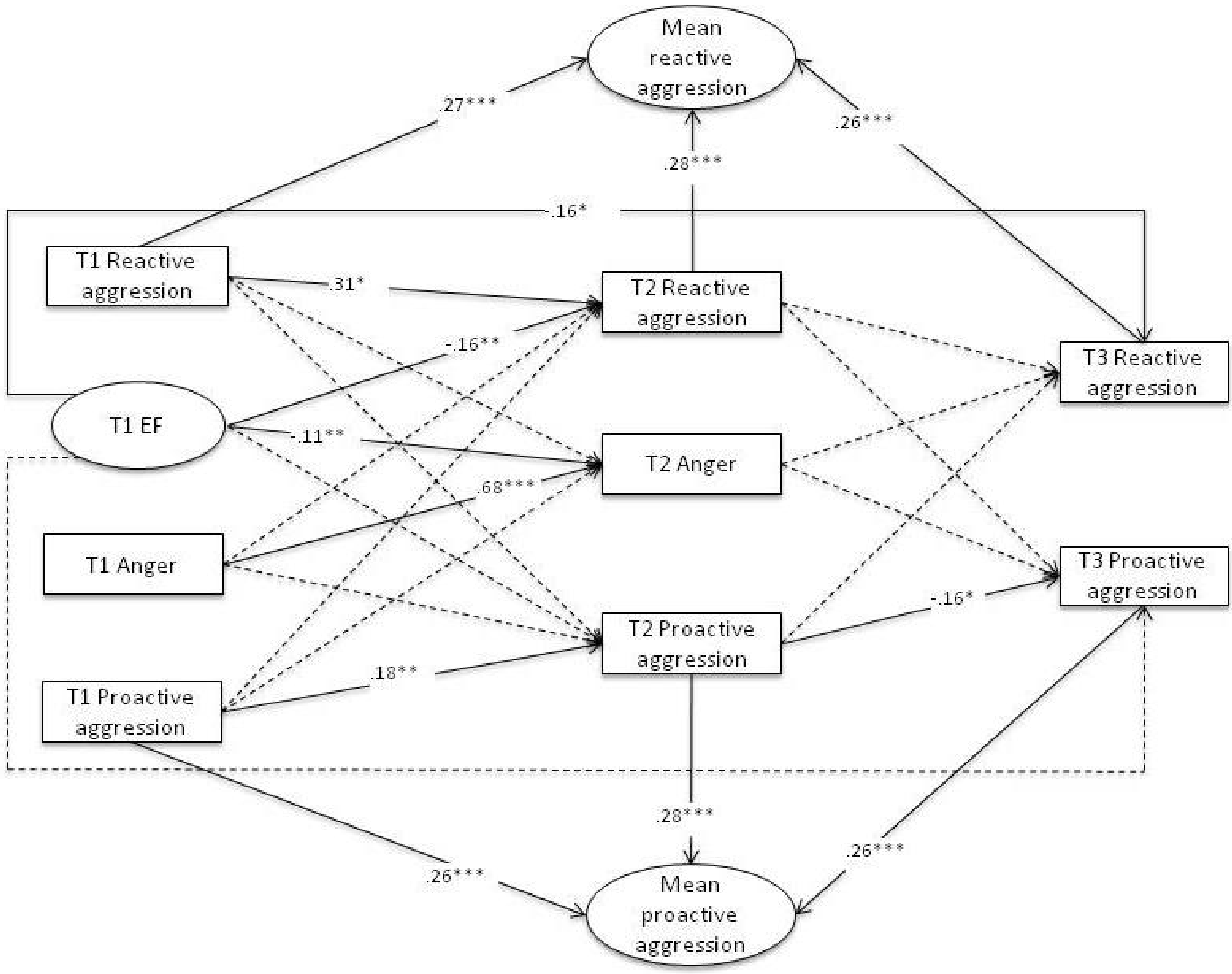
Study 2

Figure 1



Study 2

Figure 2



5. Study 3

Holl, A. K., Vetter, N. C., & Elsner, B. (2018). Disentangling the effect of theory of mind and executive function on conduct problems in middle childhood and early adolescence. Manuscript submitted for publication.

Study 3

Abstract

Conduct problems are associated with cognitive deficits in theory of mind (ToM) and executive function (EF), which are positively related during childhood. However, only few studies have considered the impact of ToM and EF on conduct problems in tandem in this age-range, disentangling how they are related to conduct problems. This study examined direct and indirect cross-sectional relations between EF, ToM, and conduct problems in middle childhood and early adolescence, using a large community sample ($N = 1501$; 9-13 years old). Behavioral measures were obtained for EF, including working memory, inhibition, and set-shifting, as well as for affective and cognitive ToM, and conduct problems were assessed via parent ratings on items of the Child Behavior Checklist. Structural equation modeling revealed a positive relation of EF and ToM, a direct negative path from ToM to conduct problems, and an indirect negative path from EF via ToM to conduct problems. These paths were not moderated by gender or age-group. Thus, the study reveals differential effects of problems in understanding others' mental states and/or in regulating one's own behavior on conduct problems at the transition to adolescence. In interventions, preventing conduct problems EF and ToM should be addressed in tandem.

Study 3

Introduction

Conduct problems in middle childhood and early adolescence place children at risk for a variety of later problematic psychosocial outcomes, for instance poor mental and general health (Bevilacqua, Hale, Barker, & Viner, 2017). Aggressive behavior is an important symptom within conduct problems, among other problems like destruction of property, deceitfulness or theft, and serious violation of rules (Drabick, Steinberg, & Hampton, 2015). Different cognitive deficits have been suggested to be responsible for conduct problems including deficits in social information processing (e.g., Crick & Dodge, 1994; Harvey, 2001), in theory of mind (ToM) (e.g., Hughes & Ensor, 2006), and in executive function (EF) (e.g., Hughes, Dunn, & White, 1998).

Theory of Mind and Conduct Problems

Aggressive behavior in children with conduct problems can at least partially be explained in terms of deficits and deviance in social information processing, in particular by a bias in attributing hostile intentions to others (e.g., Dodge, 1993; Blair, Leibenluft, & Pine, 2014). Another theoretical account states that children with conduct problems display delayed or deviant understanding of others' emotions (Arsenio & Fleiss, 1996) and intentions (Dodge & Frame, 1982; Hughes et al., 1998), which in turn is linked to cognitive and affective ToM (Premack & Woodruff, 1978). According to Shamay-Tsoory and Aharon-Peretz (2007), cognitive ToM is defined as making inferences regarding other persons' beliefs, and affective ToM regarding others' emotions, respectively. Previous research has confirmed a link between deficits in social information processing, for instance in the encoding and interpretation of situational cues in children, and early deficits in ToM (Choe, Lane, Grabell, & Olson, 2013). Deficits in ToM are also proposed to hinder an important inhibiting mechanism of violence and antisocial behavior, because ToM is a prerequisite to empathic responding, and empathy and ToM are highly associated (e.g., Sharp, 2008). Empathy is critical for socialization, and deficits in empathy are manifested in children who show higher levels of conduct problems, for instance, in an impaired ability to recognize distress in peers or in lower levels of altruism toward the victim (Blair et al., 2014; Marsh et al., 2008). Further, to take another's perspective, an aspect of ToM, is important for intentional moral conduct (e.g., Hoffman, 1987). Therefore, it is likely that deficits in the ability to understand another's cognitive and affective perspectives are implicated in conduct problems (e.g., Anastassiou-Hadjicharalambous & Warden, 2008; Hare, 1970). These accounts are also referred to as the amygdala theory of violent behavior, because ToM performance is correlated with amygdala activation, and deviant functioning of the amygdala has been found in youth with conduct disorder (e.g., Blair, 2005; Blair et al., 2014; Sebastian et al., 2012; Sharp, 2008).

However, results on the relation between deficits in ToM and aggressive or conduct problems during childhood are mixed. Hard-to-manage preschoolers performed worse in ToM tasks in comparison to controls (Hughes et al., 1998), however, their performance in cognitive and affective ToM did not correlate with their conduct problems (Hughes, White, Sharpen, & Dunn, 2000). Similarly, 6- to 12-year-old children with conduct disorder mastered two simple cognitive ToM task, but showed impaired ToM ability in items relevant for everyday life (Happé & Frith, 1996). As a consequence, children with conduct problems may rather have – instead of an impaired – an intact but deviant ToM, or a 'theory of nasty minds' (Happé & Frith, 1996, p. 395). So, children with conduct problems may use their intact ToM for antisocial behavior like lying, cheating, or teasing others (Happé & Frith, 1996). Similarly, 7- to 10-year-old ringleader bullies displayed superior cognitive and affective ToM compared to follower bullies (Sutton, Smith, & Swettenham, 1999).

Different reasons might be responsible for the mixed results: First, they could be a consequence of applying too easy ToM tasks that do not assess abilities that are relevant to real-life social interactions (Richell et al., 2003). In line with this, in more advanced tests, like "reading the emotions of the eyes", an affective ToM task, 9- to 13-year-olds with conduct problems displayed impaired ToM ability (Sharp, 2008). A second reason may be the failure to take the affective and cognitive facet of ToM into account in age-appropriate tasks, because children with conduct disorder displayed distinct deficits in both facets in middle (Anastassiou-Hadjicharalambous & Warden, 2008) and early childhood (de la Osa, Granero, Domenech, Shamay-Tsoory, & Ezpeleta, 2016). Third, not all studies controlled ToM for verbal abilities, an important correlate of ToM development (Milligan, Astington, & Dack, 2007). And finally, another unresolved issue is the potential moderation of the association between ToM and conduct problems by gender or age. Gender is a potential moderator because some studies have observed superior ToM ability in girls compared to boys (e.g., Bosacki & Astington, 1999; Calero, Salles, Semelman, & Sigman, 2013; Devine & Hughes, 2013), and boys are reported to show more conduct problems than girls (e.g., Ravens-Sieberer et al., 2008). When it comes to age as a potential moderator, it is important to note that conduct problems increase with the beginning of adolescence, where the gender difference in prevalence rates among females and males with conduct problems is narrowest (e.g., Moffitt, Caspi, Dickson, Silva, & Stanton, 1996). Moreover, ongoing development of ToM abilities was confirmed by neural and behavioral evidence throughout middle childhood (e.g., Apperly, Warren, Andrews, Grant, & Todd, 2011; Banerjee, Watling, & Caputi, 2011; Devine & Hughes, 2013; Lecce, Bianco, Devine, & Hughes, 2017) and adolescence (Vetter, Leipold, Kliegel, Phillips, & Altgassen, 2013), parallel to major changes in children's peer relations with the transition to adolescence (Steinberg & Morris, 2001). Because of these limitations of some of the previous research, the present study included ToM tasks that assess the cognitive and affective facet of ToM with sufficient difficulty to detect deficits in children with conduct problems, and one assesses abilities that are relevant for everyday social interaction. Further, we controlled ToM for verbal ability, and explored potential moderation by gender and age.

Study 3

Executive Function and Conduct Problems

EF is an umbrella term that is usually equated with the set of higher-order processes associated with the prefrontal cortex that govern goal-directed action and adaptive responses to novel, complex, or ambiguous situations (Hughes, Graham, & Grayson, 2005). Assessment of EF is difficult, because every EF task assesses systematic variance attributable to non-EF processes (Miyake & Friedman, 2012). Using structural equation modeling and a latent-variable framework solves this problem of 'task impurity', and identified three EF factors (Miyake et al., 2000): Working-memory updating refers to the ability to maintain and manipulate information in working memory for short time-intervals (Huizinga, Dolan, & van der Molen, 2006). Inhibition is the ability to withhold or restrain from a prepotent motor-response, and mental set-shifting is the ability to actively switch between different mental rule-sets or tasks (Miyake et al., 2000). The 3-factor conceptualization of EF is widely accepted and was confirmed in studies with children after the age of six (e.g., Lehto, Juujarvi, Kooistra, & Pulkkinen, 2003; Huizinga, Dolan, & van der Molen, 2006). Paralleling the protracted maturation of prefrontal cortex, EF continues to develop throughout adolescence (Blakemore & Choudhury, 2006).

Blair (2016) considers EF as one component of a larger system of self-regulation, and hypothesizes that EF builds the foundation for intentional, self-directed control of behavior. Fittingly, aggressive behavior of children with conduct problems has also been interpreted in terms of impairment in behavioral control (Hughes et al., 1998) and in solving social problems (Séguin & Zelazo, 2005), which is both attributed to deficits in EF. Because of the complexity of social situations, it is cognitively demanding to accurately process all relevant information, and thus, social-information processing might also be impaired in children with deficits in their EF (Zadeh, Im-Bolter, & Cohen, 2007). Similarly, the frontal-lobe hypothesis of emotional and behavioral regulation proposes that cognitive function of the prefrontal cortex is systematically impaired in individuals who show aggressive behavior (Moffitt, 1993;

Séguin, 2009). Consequently, deficits in EF are another important socio-cognitive risk factor for conduct problems in children (e.g., Nigg, 2000). The association between deficits in EF and aggressive behavior has been confirmed by a large body of research in early childhood, for instance in a meta-analysis among preschoolers (Schoemaker, Mulder, Deković, & Matthys, 2013; Poland, Monks, & Tsermentseli, 2016), in middle childhood (e.g., Eisenberg et al., 2004; Huang-Pollock, Shapiro, Galloway-Long, & Weigard, 2017), and also across the lifespan (Ogilvie, Stewart, Chan, & Shum, 2011). However, in adolescence, only one recent study has explored this link and has found significant positive correlations between error scores of EF tasks and conduct problems (Dolan & Lennox, 2013). Older evidence comes from studies mainly including adolescents that were incarcerated or listed in criminal records (for a review, see Morgan & Lilienfeld, 2000).

Regarding gender differences in the relation between EF and conduct problems, evidence is scarce. This is surprising because of higher conduct scores in boys (e.g., Ravens-Sieberer et al., 2008) and preliminary evidence on girls' superiority in EF tasks (e.g., Berlin & Bohlin, 2002; Carlson & Moses, 2001). Further, in children who displayed conduct problems, boys showed higher impairment in EF than girls (Lonigan et al., 2017; Raaijmakers et al., 2008). Evidence on potential moderation by age in this link is also scarce, but Schoemaker et al. (2013) found a stronger relation in older as compared to younger preschoolers. Investigation of age differences deems important because of the protracted development of EF in middle childhood and adolescence (e.g., Blakemore & Choudhury, 2006). Altogether, following previous evidence, we assessed the three facets of EF, working memory updating, inhibition, and set-shifting, to examine the relation between EF and conduct problems in middle childhood and early adolescence. Further, we explored moderation by gender and age.

Theory of mind, Executive Function and Conduct Problems

ToM and EF are often studied in tandem, because of their relative interdependence. A recent meta-analysis reported a correlation of *r* = .38 in 3- to 6-year-old children from 15 countries

(Devine & Hughes, 2014). Considering longitudinal relations in middle childhood, evidence so far supports the idea that earlier EF predicts later ToM (Austin, Groppe, & Elsner, 2014; Lecce, Bianco, Devine, & Hughes, 2017). This is in line with *executive performance* and *competence* accounts, which state that EF is a requirement for children to express their ToM abilities (e.g., Apperly, Warren, Andrews, Grant, & Todd, 2011). For instance, children might need some inhibitory ability to distance themselves from their own point of view, or cognitive flexibility to shift their perspective away from their own point of view to another's. However, research considering ToM and EF as predictors of conduct problems in tandem in middle childhood and early adolescence is still limited. There is some research with younger children, but this has yielded mixed evidence on the cross-sectional links between ToM, EF and conduct problems. For instance, hard-to-manage preschoolers scored significantly below a matched control group on a cognitive and an affective ToM task and an EF task, but EF was associated with ToM only in the hard-to-manage group (Hughes et al., 1998). This suggests a direct link between deficits in EF and conduct problems, but also an indirect link through impaired ToM (Hughes et al., 1998). Likewise, some studies in preschool-age provide evidence for direct effects of deficits in EF, independent of deficits in ToM, on conduct problems (e.g., O'Toole, Monks, & Tsermentseli, 2017), but others do not (e.g., Hughes & Ensor, 2006). In 6- to 7-year-old children referred for attention and behavior problems, deficits in EF and ToM were related to social problems (Fahie & Symons, 2003). However, this study used a measurement framework for EF that is different to the one proposed by Miyake et al. (2000). Regarding gender differences, the relations between EF, ToM, and aggressive behaviors in 3- to 6-year-olds were not moderated by gender (O'Toole et al., 2016).

Two longitudinal studies provide preliminary evidence, namely, deficits in 3-year-old children's effortful control, a cognitive construct that is closely related to EF (Bridgett, Oddi, Laake, Murdock, & Bachmann, 2013), predicted children's concurrent and later peer aggression at age 6 (Olson, Sameroff, Kerr, Lopez, & Wellman, 2005; Olson, Lopez-Duran, Lunkenheimer, Chang, & Sameroff, 2011). However, children's deficits in ToM predicted only concurrent, but not later peer aggression. Moreover, in these studies, the unique influences of ToM and effortful control were examined in separate regression analyses.

In sum, children with conduct problems show deficits in processing social information (e.g., Dodge, 1993) and related deficits in ToM (e.g., Hughes et al., 1998), which can lead to incorrect interpretation of social situations, for instance based on others' hostile intents, and/or to impairment in the violence inhibiting mechanism (e.g., Sharp, 2008; Blair et al., 2014). Further, children who display conduct problems and higher levels of aggressive behavior tend to also show related deficits in processing social information, impaired social problem-solving and behavioral control, which is attributed to deficits in EF (e.g., Zadeh et al., 2007; Séguin & Zelazo, 2005). Moreover, EF and ToM are interrelated throughout development (Devine & Hughes, 2014), and previous research in younger children suggests direct links between deficits in EF and conduct problems as well as indirect links through impaired ToM abilities (Hughes et al., 1998). However, there is mixed evidence on the effect of deficits in ToM and EF on conduct problems, when they are studied in tandem and in one analysis, and research on this issue is limited particularly in middle childhood and early adolescence. In addition, potential moderation by gender and age in these relations was rarely investigated.

Our Study

Our aim was to study direct and indirect effects of ToM and EF in tandem and in one analysis, and thereby to disentangle their respective impact on conduct problems in middle childhood and early adolescence. This was achieved by the application of structural equation modeling, using a latent-variable approach. We assessed behavioral measures of cognitive and affective ToM, as well as of the EF indicators updating, set-shifting, and inhibition in a large community sample. Conduct problems were rated by children's parents.

Study 3

Based on theoretical considerations and previous evidence outlined above, we postulated three hypotheses: First, we expected a negative path from ToM to conduct problems (e.g., Sharp, 2008), such that low ToM would predict higher levels of conduct problems (Hypothesis 1). Second, we expected a positive path from EF to ToM (Devine & Hughes, 2014) (Hypothesis 2), thereby following the accounts that assume a dependence of ToM deployment on EF competence or performance (e.g., Apperly et al., 2011). Thus, we proposed that children and young adolescents with low EF would also show low ToM abilities. And third, we hypothesized a direct negative path from EF to conduct problems (Hypothesis 3a) (Ogilvie et al., 2011). Further, we examined whether a negative indirect effect from EF to conduct problems through ToM is present (Hypothesis 3b) (Hughes et al., 1998). In addition, we tested potential moderation by gender in the postulated paths due to gender differences in the prevalence of conduct problems (e.g., Ravens-Sieberer et al., 2008), female superiority in ToM (e.g., Devine & Hughes, 2013), and preliminary evidence on gender differences in the link between EF and conduct problems (Raaijmakers et al., 2008). Finally, we explored moderation in our model's relations by age-group (3 groups, each containing 1/3 of the children) in light of the protracted development of some aspects of ToM and EF throughout middle childhood and early adolescence (e.g., Blakemore & Choudhury, 2006) and age-related differences in the prevalence of conduct problems (Moffitt et al, 1996).

Method

Participants

The sample included 1,501 children and (young) adolescents (51.8% girls) between 9.12 and 13.76 years, with a mean age of M = 11.07 years (SD = 0.92). The study was part of a larger research project on intrapersonal developmental risk factors in childhood and adolescence from a longitudinal perspective (XXX). All children who took part in the present data wave were included in our analyses, and missing data were handled by the Full Information

Maximum Likelihood procedure (Enders & Bandalos, 2001). The actual number of observations for each study variable is noted in Table 1.

Children were recruited at 114 community primary schools representing a variety of rural and urban areas. A proxy for SES was assessed by self-reported professions (Blossfeld, 1989), which were obtained from about 80% of the mothers and 75% of the fathers. The professions were coded in a ordinal scale ranging from 0 to 4. For the returned answers, holding a job at level 4 (e.g., manager or engineer) was reported by about 13% of mothers and 19% of fathers, at level 3 (e.g., technician or nurse) by about 44% and 27%, at level 2 (e.g., carpenter or qualified office employee) by about 1% and 11%, at level 1 (e.g., salesperson or cleaning staff) by about 12% and 14%, and at level 0 (i.e., unemployed or holding a job that could not be classified) by about 11% and 5%. And about 20% of mothers and about 25% of fathers did not answer the question about their profession.

Material and Procedure

Data collection took place at the children's schools in a private room with one trained research assistant or PhD student. Parents of participating children were asked to fill in a questionnaire for each child, which they could either return by mail or fill in online. All procedures were approved by the Ethics Committee at the University of XXX and by the Ministry of Education, Youth and Sport of the Federal State of XXX. Informed consent was obtained from the primary caregiver, and the children gave verbal consent.

Conduct Problems. We used 12 items of the Diagnostic and Statistical Manual-oriented scale "Conduct Problems" of the German version of the revised Child Behavior Checklist (CBCL) (Döpfner, Plück, & Kinnen, 2014; e.g., "threatens people"; see the Results section for a list of all items). Parents were asked to indicate their child's behavior on a three-point likert scale (1 = "not true", 2 = "somewhat or sometimes true", 3 = "very true or often true"). Internal consistency was assured by Cronbach's Alpha ($\alpha = .78$). For our final model, we computed a confirmatory factor analysis (CFA) (see Results).

Study 3

Theory of Mind (ToM). We used the sum scores that equaled the number of correct responses in the affective and cognitive ToM task as indices for our latent ToM factor (see Figure 2).

Cognitive Theory of Mind was measured with a belief-desire reasoning task (Apperly et al., 2011) that was translated into German and displayed on a laptop by the experimental software E-Prime (Version 2.0, Psychology Software Tools, Inc.). Children were introduced to a boy and girl named Robert and Suzy, and they were told that they would see pictures and hear sentences (3 sentences per trial) spoken by a male and female voice, respectively. The first picture showed a table with a red and a green box, and children were informed that these boxes could contain different kind of foods (e.g., peas, cookies, mushrooms etc.). Then, children heard the first sentence, in which Susi described the reality about the position of a food item, which was not known to Robert (e.g., “The apples are in the green box”). The following picture showed Robert sitting in front of the table with the red and green box, and the children heard the second and third sentences, which described Robert’s belief (e.g., “I think the apples are in the green box”) and his desire (e.g., “I don’t like apples”). The order of sentences two and three changed in half of the trials. After hearing the three sentences, children had to indicate whether Robert would open the green or the red box. Children were told that if he liked the food he would open the box that he thought contained the food, whereas if he disliked the food he would open the other box. Children were instructed to answer as quickly and correct as possible by pressing one of two marked keys on a QWERTZ keyboard, when a black square appeared around the table. If they did not respond within a time-interval of 10 s, the next trial started automatically. In half of the trials, Robert had a true belief (B+), and in the other half he had a false belief (B-). In half of the sentences Robert expressed that he liked the food (D+), and in the other half, he disliked it (D-). This resulted in four experimental conditions, B+D+, B+D-, B-D+, B-D-, with three trials each, leading to a total of 12 trials. Across trials, the positions of the red and green box changed randomly, and also did the identity and announced hiding place of the food item. Due to time constraints, we left out one trial in each condition of the original 16 trials, which were only other randomizations of the boxes' positions and the hiding place of the food item. In two practice trials, the procedure was explained to the children, and they received feedback. In the 12 test trials, children received no feedback. Trials were displayed in a pseudorandom order, so that not two similar conditions followed each other. The dependent variable was a sum score of correct answers, with *1* for a correct, and *0* for an incorrect response, resulting in a possible range of 0 to 12. Internal consistency was good with Cronbach's ordinal Alpha, which is appropriate for categorical items, $\alpha = .88$.

Affective ToM was assessed with an adapted version of the 'facial scale' of the Cambridge Mindreading Face-Voice Battery (Golan, Baron-Cohen, & Hill, 2006), which was adapted and translated into German by Vetter and colleagues (Vetter et al., 2013; Vetter, Weigelt, Döhnel, Smolka, & Kliegel, 2014). The silent clips of 5.5 s length were displayed on a laptop by the experimental software E-Prime (Version 2.0, Psychology Software Tools, Inc.). Actors of different age groups (adolescents, young and middle-aged adults) and both gender expressed complex emotions, for instance "sullen" (see Vetter et al., 2014), in their face and torso (from shoulders upward). Gender and age of actors were equally distributed and randomized over trials. After the film clip, a choice screen presented four numbered adjectives that expressed higher affective mental-states of the actors. Children were instructed to press as quickly as possible one of four marked keys on a QWERTZ keyboard in order to choose one of the adjectives that described best how the person in the clip feels. If children did not respond within a time-interval of 10 s, the next trial started automatically. Each child received two practice trials. The order of test trials was randomized across children. The key for the correct answer was also randomized and counter balanced across trials. Every correct mental-state was used only once within the task. After a pilot-study with a sample of corresponding age, 20 film clips of the 48 of Vetter et al.'s (2014) version were chosen for average difficulty and understandability of the adjectives, describing the higher affective mental-states. Two trials were excluded from further analyses after an initial data-analysis, due to ceiling effects (90% or more correct responses). This resulted in a total of 18 items. The dependent variable was a sum score of correct answers, with *1* for a correct, and *0* for an incorrect response, resulting in a possible range of 0 to 18. Internal consistency was acceptable with Cronbach's ordinal Alpha, which is appropriate for categorical items, $\alpha = .66$.

Behavioral Measures of Executive Function. Three behavioral tasks were integrated into one latent factor to measure behaviorally assessed EF, according to the model proposed by Miyake et al. (2000).

Set Shifting was assessed with a Dimensional Change Card Sorting task (Qu, Low, Chong, Lim, & Keren-Happuch, 2013) that we adapted, translated into German, and presented on a small laptop with the software E-Prime (Version 2.0). Children learnt across several trials to sort stimulus figures according to one rule, which was either "shape" or "color". One geometrical form (see Figure 1) that varied along two dimensions (shape or color) across trials was presented centrally at the lower part of the screen. Children had to press the left or the right one of two marked keys on a QWERTZ keyboard to sort this stimulus to one of two figures presented above that target stimulus – that had either similar shape or color – according to the rule that was displayed by a white-lettered word in the center of the screen. In the first 20 trials (baseline), children always had to sort after the same rule (counterbalanced across children: "shape" or "color"). After a short break, the experimenter explained that the rule could change now and that the child would have to be alert to sort after the according rule. Of the following 48 test trials, 36 were non-switch trials (sorted after the baseline rule) and 12 were switch trials (the rule changed without forewarning). The order of the 48 test trials and the mapping of keys and stimulus dimensions were randomized across children. Children were introduced to the task and the two sorting rules in two practice trials, and then told to answer as quickly and correct as possible. The dependent variable was the proportion of correct switch trials in relation to the total number of switch trials, which resulted in a score with a possible range of 0 to 1.

Inhibition was assessed by a Fruit Stroop Task (Röthlisberger, Neuenschwander, Michel, & Roebers, 2010) originally developed by Archibald and Kerns (1999), in which 4 pages with 25 stimuli each were presented to the child consecutively [Page 1: colored rectangles; Page 2: four kinds of fruits/ vegetables in correct colors; Page 3: the same fruits/ vegetables, but only printed in grey; Page 4: same fruits/ vegetables, but colored incorrectly]. Children were instructed to name the color in which the items were printed (page 1 and 2) or the color that the fruits and vegetables should have (page 3 and 4), as fast and correct as possible by saying it loudly. Response times (RT) that children needed to correctly name all colors on every page were measured, resulting in an interference score with higher values indicating more interference: RT p.4 - [(RT p.1 × RT p.3)/(RT p.1 + RT p.3)] (Röthlisberger et al., 2010).

Working Memory Updating was tested with the Digit-Span Backwards Task of the German version of the Wechsler Intelligence Scale for Children (Petermann & Petermann, 2007). Children had to verbally repeat sequences of numbers in reverse order after hearing them from the experimenter. Each trial consisted of two sequences of numbers. The first and second were 2 digits long, in each of the following trials the sequences increased by one digit up to a maximum of 8 digits. After two wrong answers in one trial, the task was ended. The dependent variable was a sum score of correctly repeated sequences, with a potential range of 0 to 14.

Information processing capacity. We assessed children's information processing capacity with the Digit-Symbol Test (DST) of the German version of the Wechsler Intelligence Scale for Children (Petermann & Petermann, 2007). Children had to assign symbols to 9 digits according to a key, which was explained before and visible throughout the task, as fast as possible. The dependent variable was the number of correctly allocated symbols within 120 seconds, expressed as standardized T-values.

Verbal ability. We controlled the ToM latent factor scores for verbal ability, because our ToM measures required good verbal understanding. We used a paper and pencil-task (Weiß, 2010) that measures the colloquial vocabulary above and beyond the German core vocabulary and offers optimal discrimination from grade 3 to 6 in German elementary schools. In 30 items, children had to choose the correct one of 5 synonyms for a given test word by marking it on a worksheet. The dependent variable was a sum score of correct answers, with *1* for a correct, and *0* for an incorrect response, resulting in a possible range of 0 to 30.

Statistical analyses

We computed structural equation models to test our hypotheses using M*plus* (Version 7.4; Muthén & Muthén, 2012) and conducted descriptive analyses with SPSS (Version 25). In all models, we used the robust Maximum Likelihood estimator (MLR) to account for non-normal distribution of our data. Age and gender were covariates for all constructs, information processing capacity was a covariate for EF, and verbal ability was a covariate for ToM. Additionally, we computed separate multi-group analyses to test for potential moderating effects of gender and age-groups in our model. Age-groups (‘younger’ vs. ‘middle’ vs. ‘older’) were defined on splitting the sample by age at the 33.33th and 66.66th percentile ($Md_{33\%}$ = 10.57 years, $Md_{66\%}$ = 11.60 years). The measurement invariance between gender and age-groups was assessed by comparing a model in which all paths were restricted to be equal (fully constrained) with a model in which all paths were allowed to vary between groups (fully unconstrained). The indicator for measurement invariance was a nonsignificant difference in χ^2 with scale corrections for the MLR estimator, as proposed by Satorra and Bentler (2001), or a nonsignificant Wald test for invariance in specific paths and indirect effects.

The percentage of missing values in our data ranged from 29.8 % in the parent-rated conduct problems to 0.01% in child data for updating (for the sample size of each measure see Table 1). Missing data in the structural equation models were handled by Full Information Maximum Likelihood (FIML) estimation in order to avoid a reduction in sample size. This procedure is considered as best practice in handling missing data (Enders & Bandalos, 2001).

Overall model fit was evaluated using the following criteria (Brown, 2006; Hu & Bentler, 1999): a root-mean-square error of approximation (RMSEA) ≤ .06, a comparative fit index (CFI) ≥ .90, and a standardized root-mean-square residual (SRMR) ≤ .08 indicated a good fit. We reported χ^2-statistic for providing complete information, but we did not evaluate it as an absolute indicator of fit, because it tends to be biased for large sample sizes (Schermelleh-Engel, Moosbrugger, & Müller, 2003). Indirect effects were tested by bootstrap analyses. If the bootstrapped 95% confidence interval does not include zero, the indirect effect is considered to be significant (Shrout & Bolger, 2002). Effect sizes for *t*-tests are reported with the conventional interpretation of $d = .20$ as small, $d = .50$ as medium, and $d = .80$ as large effects (Cohen, 1988). For our SEM, we report standardized regression coefficients as effect sizes, which were interpreted as small from .01 to .30, as moderate from .30 to .50, and large effects were > .50 (Kline, 2011).

Results

Descriptive Statistics, Correlations, and Factor Analyses

Table 1 presents all descriptive statistics and the results of *t*-tests for gender differences in all variables. In the following, only significant results are reported. Girls exhibited higher ability in affective ToM than boys with a small effect size, but no difference in cognitive ToM occurred. Similarly, girls also displayed higher abilities in inhibition and set-shifting than boys with small effect sizes, but there was no gender difference in updating. Boys showed more parent-reported conduct-problems than girls with a moderate effect. Finally, in verbal ability no gender difference was present, but in the digit-symbol-test as a proxy for information-processing capacity girls displayed higher values than boys.

The correlations among all study variables are depicted in Table 2. ToM and EF were positively correlated in a medium range. ToM further correlated negatively with conduct problems, positively with age and information-processing capacity – all in a small range. ToM showed a high and positive correlation with verbal ability, and a small negative correlation with gender, indicating higher values for girls. EF did not correlate significantly with conduct problems, but correlated positively with information-processing capacity in a medium range. Further, EF showed a positive moderate correlation with age, a small positive correlation with verbal ability, but no correlation with gender. Conduct problems did not significantly correlate with age and information-processing capacity, but they displayed a small negative correlation with verbal ability and a small positive correlation with gender.

Prior to running the path analysis, we determined an adequate measurement model for conduct problems via a confirmatory factor analysis. As the final measurement model, we determined a hierarchical model with one second-order factor and three first-order factors, which showed an excellent fit ($\chi^2[49] = 53.75$, $p < .001$; RMSEA = .01, CI .90 [0, .02]; CFI = .99; SRMR = .03). All 12 item loadings on the three first-order factors were significant, all *p*s < .01, and so were the factor loadings on the second-order factor, $\beta_{F1} = .80$, $\beta_{F2} = .84$, $\beta_{F3} = .64$, all *p*s < .001. Four items that included aggressive behavior against others ("Cruelty, bullying, or meanness to others", "Threatens people", "Gets in many fights", "Physically attacks people") loaded on the first factor. Therefore we named this factor 'aggressive behavior'. Six items that described delinquent or deviant behavior ("Breaks rules at home, school, or elsewhere", "Sets fires", "Doesn't seem to feel guilty after misbehaving", "Hangs around with others who get in trouble", "Swearing or obscene language", "Runs away from home") loaded on the second factor which was named 'deviant behavior'. Finally, two items that included devastation of things ("Vandalism", "Destroys things belonging to himself/

herself, or his/her family, or others") loaded on the third factor, which was named 'devastation'.

Hypotheses-testing analyses

We tested our hypotheses in one structural equation model (see Figure 2) that showed an acceptable fit ($\chi^2[174] = 394.89$, $p < .001$; RMSEA = .03, CI .90 [.03, .03]; CFI = .91; SRMR = .04). Information-processing capacity was included as a covariate for EF, and verbal ability was included as a covariate for ToM, and age and gender as a covariate for all factors. The coefficient of determination for conduct problems was $R^2 = .09$ in our model. Our first hypothesis was confirmed with a significant negative path from ToM to conduct problems while controlling for effects of EF, $\beta = -.23$, $p < .001$. Thus, the lower the child's ToM score was, the higher was the parent-rated conduct-problems score. The second hypothesis was also confirmed with a significant positive path from EF to ToM, $\beta = .26$, $p < .001$, which means that children who had lower EF skills also had lower ToM scores. Considering our third hypothesis, the direct path from EF to conduct problems was not significant, $\beta = .03$, $p = .644$. However regarding hypothesis 3b, we found a significant negative indirect effect from EF via ToM to conduct problems, $\beta = -.06$, $p = .002$, 95% CI [-.10, -.03]. Thus, children's lower EF scores were only associated to higher conduct problems through lower ToM abilities.

Multi-Group Analyses

To examine potential gender differences in our model, we compared a fully unconstrained model (fit: $\chi^2[314] = 500.13$, $p < .001$; RMSEA = .03, 90% CI [.02, .03]; CFI = .92; SRMR = .05), in which all paths were allowed to vary between girls and boys, with a fully constrained model (fit: $\chi^2[356] = 659.82$, $p < .001$; RMSEA = .03, 90% CI [.03, .04]; CFI = .87; SRMR = .11), in which all paths were constrained to be equal between girls and boys. The difference in χ^2 between these two models was significant, $\Delta\chi^2(42) = 117.07$, $p < .001$, indicating gender differences in specific parts of the model. Therefore we computed a revised model, in which we allowed those factor loadings, intercepts, and covariances to vary freely that were found to differ between girls and boys (e.g., two intercepts of the conduct problems items). This revised model showed an acceptable fit (χ^2[349] = 522.65, $p < .001$; RMSEA = .03, 90% CI [.02, .03]; CFI = .92; SRMR = .07), and fitted significantly better than the fully constrained model, $\Delta\chi^2(7) = 59.64$, $p < .001$, but did not fit significantly worse than the fully unconstrained model, $\Delta\chi^2(35) = 36.87$, $p = .382$. We tested all paths that were relevant for our hypotheses on differences between girls and boys, with non-significant results, all $W[1] \leq 0.26$, ps $\geq .608$. Likewise, the indirect effect did also not differ significantly between girls and boys, $W[1] = 0.20$, $p = .657$.

We followed the same approach to examine potential age-group differences in our model. We first compared a fully unconstrained model (fit: χ^2[477] = 813.68, $p < .001$; RMSEA = .04, 90% CI [.03, .04]; CFI = .87; SRMR = .06), in which all paths were allowed to vary between the three age-groups, with a fully constrained model (fit: χ^2[561] = 906.25, $p < .001$; RMSEA = .04, 90% CI [.03, .04]; CFI = .86; SRMR = .10), in which all paths were constrained to be equal. The difference in χ^2 between these two models was significant, $\Delta\chi^2(84) = 113.31$, $p = .018$, indicating age-group differences in specific parts of the model. Therefore we computed a revised model, in which we allowed those intercepts and one variance that were found to differ between two of the three groups to vary freely in the older group (e.g., one intercept of the conduct problems items). This revised model showed an acceptable fit (χ^2[558] = 880.46, $p < .001$; RMSEA = .03, 90% CI [.03, .04]; CFI = .87; SRMR = .09), and fitted significantly better than the fully constrained model, $\Delta\chi^2(3) = 12.35$, $p = .006$, but did not fit significantly worse than the fully unconstrained model, $\Delta\chi^2(81) = 95.37$, $p = .131$. We tested all paths that were relevant for our hypotheses on differences between all three age-groups, with non-significant results, all $W[1] \leq 0.91$, ps $\geq .34$. Likewise, the indirect effect did not differ significantly between the three age-groups, all $W[1] \leq 0.56$, ps $\geq .455$.

Study 3

Discussion

In this cross-sectional study, we disentangled the effect of ToM and EF on conduct problems in a large community sample with an age range from middle childhood to early adolescence (9-13 years). ToM and EF were both assessed experimentally in a multi-variable approach via different tasks for affective and cognitive ToM, and for the EF-facets updating, inhibition, and set shifting, respectively. Conduct problems were rated by children's parents on 12 items of the CBCL conduct problems scale. In our structural equation model, we controlled all variables for age and gender, as well as ToM for verbal ability and EF for information-processing capacity. In addition, we tested potential moderation by gender and age-group (3 groups, each containing 1/3 of the children) in our model via multi-group-models.

Our first hypothesis was confirmed by a significant negative path from ToM to conduct problems. Thus, children in our study who displayed lower ToM abilities (with respect to predicting another person's behavior based on that person's beliefs and desires as well as to identifying other persons' emotional expressions) showed more conduct problems. This finding is in line with previous research in middle childhood (e.g., Anastassiou-Hadjicharalambous & Warden, 2008) and early adolescence (Sharp, 2008). Our finding confirms that children who display conduct problems have impaired ToM, as proposed by the amygdala theory of violent behavior (e.g., Blair et al., 2014). It could be that because ToM is a prerequisite to empathy (e.g., Shamay-Tsoory & Aharon-Peretz, 2007), the mechanism that should hinder violent behavior against other persons or their property is impaired in children with conduct problems.

The result can also be explained in terms of the social information-processing theory (Crick & Dodge, 1994), which proposes that children with deficits in their processing of relevant social information show more aggressive behavior than other children. Because ToM is defined as the ability to infer others' mental states, it can be assumed to play a notable role in social information-processing, for instance in the attribution of hostile intent to others (e.g., Choe et al., 2013). As a consequence, children with deficits in their ToM could also display impaired processing of relevant social cues, which might lead them to behave in aggressive or anti-social ways against other children.

Our second hypothesis of a significant positive path from EF to ToM was also confirmed. Thus, children with higher EF abilities also showed higher ToM abilities in our study. This is in line with previous evidence on the relation between ToM and EF in middle childhood (e.g., Lecce et al., 2017) and in adolescence (e.g., Vetter et al., 2013). This result also extends previous research, because we assessed cognitive ToM as well as affective ToM, which has been neglected in previous research. Our finding could be interpreted in light of the expression or competence accounts that propose that EF is a prerequisite for the development of ToM abilities (e.g., Apperly et al., 2011). However, because our study is cross-sectional, this result cannot be interpreted as a causal effect, and the direction of effect has to be determined in future longitudinal studies.

Our third hypothesis proposed a negative direct path from EF to conduct problems, and a possible negative indirect path from EF via ToM to conduct problems. The first part of the hypothesis was not confirmed in our study, due to a non-significant direct path. However, we found a significant negative indirect path from EF via ToM to conduct problems. Consequently, children with deficits in their ability to update relevant information in their working memory, to inhibit prepotent behavioral impulses, and to shift between different rule systems also displayed impairment in inferring the affective and cognitive mental states of others, also exhibited higher levels of conduct problems. This could be explained in terms of the competence account that proposes that sufficiently developed EF is a prerequisite for the application of ToM abilities (e.g., Apperly et al., 2011). It should be noted, however, that our indirect link was only of small size, and that our study was only cross-sectional, and thus this result should be interpreted with caution.

Study 3

Our finding does not confirm the frontal lobe hypothesis of aggressive behavior that proposes that children who display conduct problems show deficits in EF (e.g., Séguin, 2009), or that their deficits in EF can be interpreted in terms of impairment in behavioral control (e.g. Hughes et al., 1998) or social problem-solving (e.g., Séguin & Zelazo, 2005). It is also not in line with previous research on a direct link between deficits in EF and conduct problems in mostly younger age-groups (e.g., Schoemaker et al., 2013) or across the life-span (Ogilvie et al., 2011). The absence of a direct link between EF and conduct problems could be explained by differences in participants' age or clinical status, different measures for conduct problems in contrast to conduct disorder, or the examination of EF and ToM in the relation to conduct problems in tandem. For example, Olson et al. (2011) investigated the effects of effortful control and ToM in separate regression analyses. By contrast, we examined both possible predictors in one structural equation model, thereby controlling the path from EF to conduct problems for its relation to ToM. However, the regression of conduct problems on EF as a sole predictor was also not significant in our study ($\beta = -.09$, $p = .11$). Moreover, other research also failed to determine a direct significant effect from EF to conduct problems in a regression analysis. For instance, Hughes and Ensor (2006) did not find an independent effect of EF on behavioral problems, and an earlier quantitative review neither found a relation between EF and conduct disorder (Pennington & Ozonoff, 1996). In addition, Morgan and Lilienfeld (2000) argued in a meta-analytic review that the specificity of EF deficits to conduct disorder is not clear, because persons who exhibit antisocial behavior also display deficits on other neuropsychological tasks. As a consequence, children with conduct disorder could display deficits in neuropsychological tasks in general, so deficits in EF could be a possible consequence, but not a sufficient proposition for conduct disorder and problems.

In additional multi-group analyses, we explored possible moderation of gender and age-group in our model. We did not find indications that the hypothesized paths differed by gender, despite boys exhibiting more conduct problems than girls in mean level differences with a moderate effect. This is in line with previous research showing that mostly, boys display more conduct problems than girls (e.g., Ravens-Sieberer et al., 2008), and is also in accordance to our finding that some intercepts of the conduct-problems factor in our multi-group model had to be estimated freely. However, in this revised model no gender differences were found in the hypothesized paths. Our finding is in contrast to two previous studies on aggressive behavior in preschool children, in which boys exhibited more EF impairment than girls, however, our study focused on another age-range (Lonigan et al., 2017; Raaijmakers et al., 2008). But other studies also did not find a moderation by gender in their model, for instance in the relation of working memory and externalizing disorders (e.g., Huang-Pollock et al., 2017). Nevertheless, we found gender differences in mean levels of EF, in that girls showed less interference in inhibition meaning higher inhibition abilities, and also displayed higher set-shifting and information-processing capacity. Research on gender differences in EF in middle childhood and early adolescence is very limited, and preliminary evidence indicates superior EF abilities in girls (e.g., Berlin & Bohlin, 2002; Carlson & Moses, 2001). However, other studies also found lower scores for boys in effortful control (Olson et al., 2005), but did not find a moderation in the relation between effortful control and concurrent and later peer aggression (Olson et al., 2011).

In addition, we also found gender differences in mean levels of affective ToM, favoring girls. Girls' superiority in ToM abilities was found in previous research (e.g., Devine & Hughes, 2013), which is mainly explained by different play behaviors in girls, who prefer pretend play, which fosters social interaction, as compared to boys, who prefer activities that involve mainly spatial abilities. Other explanations encompass biological differences between the two gender groups, for instance a general female superiority in abilities that are related to empathy (Baron-Cohen, 2002); albeit most research until now has considered only cognitive ToM, and less is known about affective ToM. We found no gender differences in cognitive ToM, which might have been a consequence of a small ceiling effect in our cognitive ToM measure. Future research should investigate gender differences in ToM in middle childhood and early adolescence with tasks that provide sufficient interindividual variance in this age, and determine whether girls also show superiority in affective ToM. In sum, the underlying mechanisms of effect from ToM on conduct problems under the control of EF seems to work in the same way for both gender groups despite mean level differences. Consequently, the two gender groups seem to share more similarities in these relations than differences, and conclusions from our study can be drawn for girls and boys similarly, which makes application of our results, for instance in interventions, easier.

As a further research aim, we also explored moderation in our model's relations by age-group, motivated by the ongoing development of both EF (Blakemore & Choudhury, 2006) and ToM abilities in middle childhood (e.g., Apperly et al., 2011) and adolescence (e.g., Vetter et al., 2013). This was investigated in additional multi-group models with three age-groups that were identified by splits on the 33.33 and 66.66 percentiles ($Md_{33\%}$ = 10.57 years, $Md_{66\%}$ = 11.60 years). Again, we did not find indication that the hypothesized paths in our model differed by age. However, we found significant positive correlations of age with ToM and EF, which indicates that ToM and EF scores were relatively higher in older children in our study. This is in line with previous research that found neural evidence on a protracted development of ToM and EF throughout middle childhood and adolescence (e.g., Blakemore & Choudhury, 2006; Vetter et al., 2013). Nevertheless, the hypothesized paths from ToM and EF to conduct problems and from EF to ToM did not alter in dependence of age. As a consequence, the underlying effects between EF, ToM, and conduct problems, for instance the violence inhibiting mechanism, seem to be similar for all three examined age-groups. This allows applying our results to the whole age-range of our sample, from 9 to 13 years, and connecting research on these relations across the age periods of middle childhood and early adolescence.

Study 3

Strengths and Limitations

Our study has several strengths. We analyzed a large community sample of school-aged children. In our model, we controlled for age and gender, but in addition, we explored moderations of gender and age-group. We also controlled ToM for verbal ability and EF for information-processing capacity, which has not been done in all previous studies. The results from our study extend previous research in several ways: Until now, associations from ToM or EF to conduct problems have often been examined separately (e.g., Olson et al., 2011). By contrast, we investigated whether deficits in ToM and EF would predict conduct problems, when these associations are examined in tandem and in one model. This allowed us to control for the correlation of ToM and EF as well as for possible overlapping variance in their relations to conduct problems, and to disentangle their effects on conduct problems. Second, we tested not only a direct effect of EF on conduct problems, but also an indirect effect from EF to conduct problems via ToM. Third, we used a latent-variable framework including not only cognitive, but also affective ToM, and three experimentally-assessed features of EF – updating, inhibition, and set-shifting. This reduced the problem of 'task impurity (Miyake & Friedman, 2012) and allowed for a broad consideration of cognitive abilities considered as decisive in this age-period (Blakemore & Choudhury, 2006). And last, we investigated the – until now – often distinctly considered developmental periods of middle childhood and early adolescence together.

We also note some limitations to our study. First, because our study was cross-sectional, causal interpretations cannot be drawn. Second, we did not investigate attention deficit and hyperactivity (ADHD) symptoms, as has been done in previous research (Antonini, Becker, Tamm, & Epstein, 2015). This could be one reason why we did not find a direct relation between EF and conduct problems, as it was argued by Pennington and Ozonoff (1996), because EF deficits could be specific to children who show conduct problems and comorbid ADHD symptoms (e.g., Antonini et al., 2015). These issues should be addressed in future studies on the relation between EF and conduct problems in middle childhood and early adolescence.

Third, we did not include assessments of hot EF (like delay discounting or decision making), which could be another reason for the absence of a direct path from EF to conduct problems. As stated by the frontal-lobe hypothesis of emotional and behavioral regulation, hot EF in particular could be impaired in children who show higher levels of aggressive behavior (Séguin, 2009). However, Poland et al. (2015) did not find a relation between hot EF and aggression in early childhood.

A fourth limitation lies in floor effects in conduct problems in our sample, and possible ceiling effects in cognitive ToM. These effects could be the consequence of our community sample that showed a rather representative distribution of different levels of possible occupations with a trend to higher level occupations. Other studies on the relation between EF, ToM, and conduct problems that focused on socially disadvantaged families found negative correlations of social disadvantage with EF and ToM, and a positive correlation with conduct problems (e.g., Hughes & Ensor, 2008). The floor and possible ceiling effects in our sample could have limited our results as a consequence of restricted variance, and may also be responsible for the missing direct effect between EF and conduct problems. A last limitation is our assessment of conduct problems. Due to time-constraints, we chose a selection of items of the DSM-oriented scale for conduct problems of the Child-Behavior-Checklist. Therefore, we cannot compare our results directly with studies that included clinical interviews or further diagnostic assessments by raters other than parents. Future studies should include multi-informant ratings and a broader range of items to assess conduct problems, for instance all items of the CBCL conduct problems scale (e.g., stealing) and also oppositional defiant problems.

Study 3

Conclusions

Despite these limitations, our findings disentangled the relations between ToM, EF, and conduct problems in an age sample that ranged from middle childhood to early adolescence. In our study, ToM negatively predicted conduct problems, EF positively predicted ToM, and EF exerted only a negative indirect effect on conduct problems through ToM. These paths were not moderated by gender or age-group. Our results highlight the necessity to consider ToM as an important negative predictor for conduct problems in middle childhood and early adolescence, which might be explained by impairment in the violence inhibiting mechanism and by deviant processing of important social cues. Regarding EF, we could not confirm proposals of a general deficit in EF in children with conduct problems, and further research is needed to investigate direct and indirect effects on conduct problems in tandem with ToM in this age-range. Interventions preventing conduct problems in the transition from middle childhood to early adolescence should foster the ability to infer others' higher affective and cognitive mental states, because ToM seems to continuously play an important role in children's socio-cognitive health. However, also EF should be on the agenda for interventions, because EF and ToM seem to stay intertwined as two important higher-order cognitive abilities in this age-period, impacting directly and indirectly on children's conduct problems.

Study 3

References

Anastassiou-Hadjicharalambous, X., & Warden, D. (2008). Cognitive and affective perspective-taking in conduct-disordered children high and low on callous-unemotional traits. *Child and Adolescent Psychiatry*, *2*(16), 1–11. http://doi.org/10.1186/1753-2000-2-16

Antonini, T. N., Becker, S. P., Tamm, L., & Epstein, J. N. (2015). Hot and cool executive functions in children with attention-deficit/hyperactivity disorder and comorbid oppositional defiant disorder. *Journal of the International Neuropsychological Society*, *21*, 584–595. http://doi.org/10.1017/S1355617715000752

Apperly, I. A., Warren, F., Andrews, B. J., Grant, J., & Todd, S. (2011). Developmental continuity in theory of mind: speed and accuracy of belief-desire reasoning in children and adults. *Child Development*, *82*, 1691–1703. http://doi.org/10.1111/j.1467-8624.2011.01635.x

Archibald, S. J., & Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychology*, *5*, 115–129. http://doi.org/10.1076/chin.5.2.115.3167

Arsenio, W. F., & Fleiss, K. (1996). Typical and behaviourally disruptive children's understanding of the emotional consequences of socio-moral events. *British Journal of Developmental Psychology*, *14*(2), 173–186. http://doi.org/10.1111/j.2044-835X.1996.tb00700.x

Austin, G., Groppe, K., & Elsner, B. (2014). The reciprocal relationship between executive function and theory of mind in middle childhood: a 1-year longitudinal perspective. *Frontiers in Psychology*, *5*(655), 1–11. http://doi.org/10.3389/fpsyg.2014.00655

Banerjee, R., Watling, D., & Caputi, M. (2011). Peer relations and the understanding of faux pas: Longitudinal evidence for bidirectional associations. *Child Development*, *82*, 1887–1905. http://doi.org/10.1111/j.1467-8624.2011.01669.x

Study 3

Baron-Cohen. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences*, *6*(6), 248–254. http://doi.org/10.1016/S1364-6613(02)01904-6

Berlin, L., & Bohlin, G. (2002). Response inhibition, hyperactivity, and conduct problems among preschool children. *Journal of Clinical Child and Adolescent Psychology*, *31*(2), 242–251. http://doi.org/10.1207/S15374424JCCP3102_09

Blair, C. (2016). Developmental science and executive function. *Current Directions in Psychological Science*, *25*(1), 3–7. http://doi.org/10.1177/0963721415622634

Blair, R. J. R. (2005). Applying a cognitive neuroscience perspective to the disorder of psychopathy. *Development and Psychopathology*, *17*, 865–891. http://doi.org/10.1017/S0954579405050418

Blair, R. J. R., Leibenluft, E., & Pine, D. S. (2014). Conduct disorder and callous-unemotional traits in youth. *The New England Journal of Medicine*, *371*(23), 2207–2216. http://doi.org/10.1056/NEJMra1315612

Blakemore, S.-J., & Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *47*, 296–312. http://doi.org/10.1111/j.1469-7610.2006.01611.x

Blossfeld, H.-P. (1989). *Kohortendifferenzierung und Karriereprozess: Eine Längsschnittstudie über die Veränderung der Bildungs- und Berufschancen im Lebenslauf [Differences in cohorts and carrier development: A longitudinal study on changes of educational and occupational chance*. Frankfurt a. M., Germany; New York, NY: Campus.

Bosacki, S., & Astington, J. W. (1999). Theory of mind in preadolescence: Relations between social understanding and social competence. *Social Development*, *8*, 237–255. http://doi.org/10.1111/1467-9507.00093

Bridgett, D. J., Oddi, K. B., Laake, L. M., Murdock, K. W., & Bachmann, M. N. (2013). Integrating and differentiating aspects of self-regulation: effortful control, executive functioning, and links to negative affectivity. *Emotion*, *13*, 47–63. http://doi.org/10.1037/a0029536

Brown, T. A. (2006). *Confirmatory factor analysis*. New York, NY: Guilford Press.

Calero, C. I., Salles, A., Semelman, M., & Sigman, M. (2013). Age and gender dependent development of theory of mind in 6- to 8-years old children. *Frontiers in Human Neuroscience*, *7*(281), 1–7. http://doi.org/10.3389/fnhum.2013.00281

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*, 1032–1053. http://doi.org/10.1111/1467-8624.00333

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*, 2245–2256. http://doi.org/10.1037/a0032293

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin*, *115*, 74–101. http://doi.org/10.1037/0033-2909.115.1.74

de la Osa, N., Granero, R., Domenech, J. M., Shamay-Tsoory, S., & Ezpeleta, L. (2016). Cognitive and affective components of theory of mind in preschoolers with oppositional defiance disorder: Clinical evidence. *Psychiatry Research*, *241*, 128–134. http://doi.org/10.1016/j.psychres.2016.04.082

Devine, R. T., & Hughes, C. (2013). Silent films and strange stories: Theory of mind, gender, and social experiences in middle childhood. *Child Development*, *84*, 989–1003. http://doi.org/10.1111/cdev.12017

Devine, R. T., & Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: A meta-analysis. *Child Development*, *85*, 1777–1794. http://doi.org/10.1111/cdev.12237

Dodge, K. A. (1993). Social-cognitive mechanisms in the development of conduct disorder and depression. *Annual Reviews Psychology*, *44*, 559–584. http://doi.org/10.1146/annurev.ps.44.020193.003015

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

Dolan, M., & Lennox, C. (2013). Cool and hot executive function in conduct-disordered adolescents with and without co-morbid attention deficit hyperactivity disorder: relationships with externalizing behaviours. *Psychological Medicine*, *43*, 2427–2436. http://doi.org/10.1017/S0033291712003078

Drabick, D., Steinberg, E., & Hampton, A. (2015). Overview of DSM disruptive behavior disorders. In T. P. Beauchaine & S. P. Hinshaw (Eds.), *The Oxford Handbook of Externalizing Spectrum Disorders* (pp. 3–18). New York, NY: Oxford University Press. http://doi.org/10.1093/oxfordhb/9780199324675.013.16

Eisenberg, N., Spinrad, T. L., Fabes, R. A., Reiser, M., Cumberland, A., Stephanie, A., … Thompson, M. (2004). The relations of effortful control and impulsivity to children's resiliency and adjustment. *Child Development*, *75*, 25–46.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*, 430–457. http://doi.org/10.1207/s15328007sem0803_5

Fahie, C. M., & Symons, D. K. (2003). Executive functioning and theory of mind in children clinically referred for attention and behavior problems. *Journal of Applied Developmental Psychology*, *24*, 51–73. http://doi.org/10.1016/S0193-3973(03)00024-8

Fergusson, D. M., Horwood, L. J., & Ridder, E. M. (2007). Conduct and attentional problems in childhood and adolescence and later substance use, abuse and dependence: Results of a 25-year longitudinal study. *Drug and Alcohol Dependence*, *88S*, 14–26. http://doi.org/10.1016/j.drugalcdep.2006.12.011

Study 3

Golan, O., Baron-Cohen, S., & Hill, J. (2006). The Cambridge mindreading (CAM) face-voice battery: Testing complex emotion recognition in adults with and without Asperger syndrome. *Journal of Autism and Developmental Disorders*, *36*(2), 169–183. http://doi.org/10.1007/s10803-005-0057-y

Happé, F., & Frith, U. (1996). Theory of mind and social impairment in children with conduct disorder. *British Journal of Developmental Psychology*, *14*, 385–398. http://doi.org/10.1111/j.2044-835X.1996.tb00713.x

Hare, R. D. (1970). *Psychopathy: Theory and research*. Oxford, UK: John Wiley & Sons.

Harvey, R. J. et al. (2001). Social reasoning: An influence on aggression. *Clinical Psychology Review*, *21*, 459–469.

Hoffman, M. L. (1987). The contribution of empathy to justice and moral judgment. New York, NY: Cambridge University Press. (N. Eisenberg & J. Strayer, Eds.)*En N. Eisenberg Y J. Strayer, Empathy and Its Development*. Cambridge: Cambridge University Press.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. http://doi.org/10.1080/10705519909540118

Huang-Pollock, C., Shapiro, Z., Galloway-Long, H., & Weigard, A. (2017). Is poor working memory a transdiagnostic risk factor for psychopathology? *Journal of Abnormal Child Psychology*, *45*, 1477–1490. http://doi.org/10.1007/s10802-016-0219-8

Hughes, C., Dunn, J., & White, A. (1998). Trick or treat?: uneven understanding of mind and emotion and executive dysfunction in "hard-to-manage" preschoolers. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *39*(7), 981–994. http://doi.org/9804031

Hughes, C., & Ensor, R. (2006). Behavioural problems in 2-year-olds: Links with individual differences in theory of mind, executive function and harsh parenting. *Journal of Child*

Psychology and Psychiatry, *47*, 488–497. http://doi.org/10.1111/j.1469-7610.2005.01519.x

Hughes, C., & Ensor, R. (2008). Does executive function matter for preschoolers' problem behaviors? *Journal of Abnormal Child Psychology*, *36*, 1–14. http://doi.org/10.1007/s10802-007-9107-6

Hughes, C., Graham, A., & Grayson, A. (2005). Executive function in childhood: Development and disorder. In I. J. Oates (Ed.), *Cognitive development* (pp. 205–230). Milton Keynes,, UK: Open University Press.

Hughes, C., White, A., Sharpen, J., & Dunn, J. (2000). Antisocial, angry, and unsympathetic: "hard-to-manage" preschoolers' peer problems and possible cognitive influences. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *41*, 169–179. http://doi.org/10.1111/1469-7610.00558

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*(11), 2017–2036. http://doi.org/10.1016/j.neuropsychologia.2006.01.010

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.

Lecce, S., Bianco, F., Devine, R. T., & Hughes, C. (2017). Relations between theory of mind and executive function in middle childhood: A short-term longitudinal study. *Journal of Experimental Child Psychology*, *163*, 69–86. http://doi.org/10.1016/j.jecp.2017.06.011

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*, 59–80. http://doi.org/10.1348/026151003321164627

Lonigan, C. J., Spiegel, J. A., Goodrich, J. M., Morris, B. M., Osborne, C. M., Lerner, M. D., & Phillips, B. M. (2017). Does preschool self-regulation predict later behavior problems in general or specific problem behaviors? *Journal of Abnormal Child Psychology*, *45*, 1491–1502. http://doi.org/10.1007/s10802-016-0260-7

Marsh, A. A., Finger, E. C., Mitchell, D. G. V, Reid, M. E., Sims, C., Kosson, D. S., … Blair, R. J. R. (2008). Reduced amygdala response to fearful expressions in children and adolescents with callous-unemotional traits and disruptive behavior disorders. *American Journal of Psychiatry*, *165*(6), 712–720. http://doi.org/appi.ajp.2007.07071145 [pii]\r10.1176/appi.ajp.2007.07071145

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, *78*, 622–646. http://doi.org/10.1111/j.1467-8624.2007.01018.x

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. http://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognitive Psychology*, *41*, 49–100. http://doi.org/10.1006/cogp.1999.0734

Moffitt, T. E. (1993). The neuropsychology of conduct disorder. *Dev Psychopathol*, *5*. http://doi.org/10.1017/S0954579400004302

Moffitt, T. E., Caspi, A., Dickson, N., Silva, P., & Stanton, W. (1996). Childhood-onset versus adolescent-onset antisocial conduct problems in males: Natural history from ages 3 to 18 years. *Development and Psychopathology*, *8*, 399–424. Retrieved from http://journals.cambridge.org

Morgan, A. B., & Lilienfeld, S. O. (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

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide. Statistical analysis with latent* *variables* (7th ed.). Los Angeles, CA: CA: Muthén & Muthén.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, *126*, 220–246. http://doi.org/10.1037/0033-2909.126.2.220

O'Toole, S. E., Monks, C. P., & Tsermentseli, S. (2017). Executive function and theory of mind as predictors of aggressive and prosocial behavior and peer acceptance in early childhood. *Social Development*, *26*, 907–920. http://doi.org/10.1111/sode.12231

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

Olson, S. L., Sameroff, A. J., Kerr, D. C. R., Lopez, N. L., & Wellman, H. M. (2005). Developmental foundations of externalizing problems in young children: The role of effortful control. *Development and Psychopathology*, *17*, 25–45. http://doi.org/10.1017/s0954579405050029

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *37*(I), 51–87. http://doi.org/10.1111/j.1469-7610.1996.tb01380.x

Petermann, F., & Petermann, U. (Eds. . (2007). *Hamburg-Wechsler-Intelligenztest für Kinder - IV (HAWIK-IV) [Wechsler intelligence scale for children IV].* Göttingen: Hogrefe.

Poland, S. E., Monks, C. P., & Tsermentseli, S. (2016). Cool and hot executive function as predictors of aggression in early childhood: Differentiating between the function and form of aggression. *The British Journal of Developmental Psychology*, *34*, 181–197. http://doi.org/10.1111/bjdp.12122

Qu, L., Low, J. J. W., Chong, P. S., Lim, D. W. L., & Keren-Happuch, F. F. E. (2013). Task switching in English-Chinese bilinguals: A life-span approach. In Y. Kashima, E. S. Kashima, & R. Beatson (Eds.), *Steering the cultural dynamics. Selected papers from the*

2010 congress of the international association for cross-cultural psychology. (pp. 76–84). University of Melbourne, Australia.

Raaijmakers, M. A. J., Smidts, D. P., Sergeant, J. A., Maassen, G. H., Posthumus, J. A., Van Engeland, H., & Matthys, W. (2008). Executive functions in preschool children with aggressive behavior: Impairments in inhibitory control. *Journal of Abnormal Child Psychology*, *36*, 1097–1107. http://doi.org/10.1007/s10802-008-9235-7

Ravens-Sieberer, U., Wille, N., Erhart, M., Bettge, S., Wittchen, H.-U., Rothenberger, A., … Döpfner, M. (2008). Prevalence of mental health problems among children and adolescents in Germany: results of the BELLA study within the National Health Interview and Examination Survey. *European Child & Adolescent Psychiatry*, *17*(S1), 22–33. http://doi.org/10.1007/s00787-008-1003-2

Richell, R. A., Mitchell, D. G. V, Newman, C., Leonard, A., Baron-Cohen, S., & Blair, R. J. R. (2003). Theory of mind and psychopathy. *Neuropsychologia*, *41*(5), 523–526. http://doi.org/10.1016/S0028-3932(02)00175-6 T4 - Can psychopathic individuals read the “language of the eyes”? M4 - Citavi

Röthlisberger, M., Neuenschwander, R., Michel, E., & Roebers, C. M. (2010). Exekutive Funktionen: Zugrundeliegende kognitive Prozesse und deren Korrelate bei Kindern im späten Vorschulalter [Executive functions: Underlying cognitive processes and their correlates in late preschool children]. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie*, *42*, 99–110. http://doi.org/10.1026/0049-8637/a000010

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514. http://doi.org/10.1007/BF02296192

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research - Online*, *8*(2), 23–74. Retrieved from https://www.dgps.de/fachgruppen/methoden/mpr-online/

Study 3

Schoemaker, K., Mulder, H., Deković, M., & Matthys, W. (2013). Executive functions in preschool children with externalizing behavior problems: A meta-analysis. *Journal of Abnormal Child Psychology*, *41*, 457–471. http://doi.org/10.1007/s10802-012-9684-x

Sebastian, C. L., McCrory, E. J. P., Cecil, C. a M., Lockwood, P. L., De Brito, S. a, Fontaine, N. M. G., & Viding, E. (2012). Neural responses to affective and cognitive theory of mind in children with conduct problems and varying levels of callous-unemotional traits. *Archives of General Psychiatry*, *69*(8), 814–22. http://doi.org/10.1001/archgenpsychiatry.2011.2070

Séguin, J. R. (2009). The frontal lobe and aggression. *European Journal of Developmental Psychology*, *6*, 100–119. http://doi.org/10.1080/17405620701669871

Séguin, J. R., & Zelazo, P. D. (2005). Executive Function in early physical aggression. In R. E. Tremblay, W. W. Hartup, & J. Archer (Eds.), *Developmental origins of aggression* (pp. 307–329). New York: Guilford Press.

Shamay-Tsoory, S. G., & Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: A lesion study. *Neuropsychologia*, *45*(13), 3054–3067. http://doi.org/10.1016/j.neuropsychologia.2007.05.021

Sharp, C. (2008). Theory of mind and conduct problems in children: Deficits in reading the “emotions of the eyes.” *Cognition and Emotion*, *22*(6), 1149–1158. http://doi.org/10.1080/02699930701667586

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*, 422–445. http://doi.org/10.1037/1082-989X.7.4.422

Steinberg, L., & Morris, A. S. (2001). Adolescent development. *Annual Review of Psychology*, *52*, 83–110. http://doi.org/0066-4308/01/0201

Sutton, J., Smith, P. K., & Swettenham, J. (1999). Social cognition and bullying: Social inadequacy or skilled manipulation? *British Journal of Developmental Psychology*, *17*(3), 435–450. http://doi.org/10.1348/026151099165384

Vetter, N. C., Altgassen, M., Phillips, L., Mahy, C. E. V, & Kliegel, M. (2013). Development of affective theory of mind across adolescence: disentangling the role of executive functions. *Developmental Neuropsychology*, *38*(2), 114–25. http://doi.org/10.1080/87565641.2012.733786

Vetter, N. C., Leipold, K., Kliegel, M., Phillips, L. H., & Altgassen, M. (2013). Ongoing development of social cognition in adolescence. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *19*, 615–629. http://doi.org/10.1080/09297049.2012.718324

Vetter, N. C., Weigelt, S., Döhnel, K., Smolka, M. N., & Kliegel, M. (2014). Ongoing neural development of affective theory of mind in adolescence. *Social Cognitive and Affective Neuroscience*, *9*(7), 1022–1029. http://doi.org/10.1093/scan/nst081

Zadeh, Y. Z., Im-Bolter, N., & Cohen, N. J. (2007). Social cognition and externalizing psychopathology: An investigation of the mediating role of language. *Journal of Abnormal Child Psychology*, *35*(2), 141–152. http://doi.org/10.1007/s10802-006-9052-9

Funding details

This work was supported by the German Research Foundation (DFG) under grant number GRK 1668/1.

Study 3

Table 1

Descriptive Statistics and Results of Statistical t-Tests for Gender Differences

	n (*n girls*)	Range	Total *M* (*SD*)	Girls *M* (*SD*)	Boys *M* (*SD*)	Gender difference
Affective ToM[a]	1479 (770)	2 - 17	10.0 (2.72)	10.30 (2.69)	9.68 (2.72)	$t(1477) = 4.41$*** $d = 0.23$
Cognitive ToM[a]	1479 (773)	1 - 12	9.55 (2.55)	9.66 (2.54)	9.42 (2.55)	$t(1477) = 1.76$
Updating[a]	1491 (775)	0 - 13	7.37 (1.63)	7.37 (1.62)	7.38 (1.65)	$t(1489) = -0.14$
Inhibition[b]	1436 (752)	1.75 - 36.72[c]	16.38 (4.74)	16.04 (4.70)	16.76 (4.76)	$t(1434) = -2.86$** $d = -0.15$
Set-Shifting[d]	1477 (771)	0 - 1	0.81 (0.16)	0.82 (0.16)	0.80 (0.15)	$t(1475) = 2.43$* $d = 0.13$
Conduct Problems[a]	1053 (548)	9 - 27	13.45 (2.14)	13.0 (1.78)	13.93 (2.38)	$t(928.21) = -7.10$*** $d = -0.44$
Verbal ability[a]	1490 (776)	2 - 30	18.98 (5.13)	18.88 (5.22)	19.08 (5.04)	$t(1488) = -0.78$
Digit-symbol-test[e]	1484 (773)	20 - 80	54.85 (8.88)	56.82 (8.98)	52.70 (8.26)	$t(1482) = 9.16$*** $d = 0.48$

Notes. ToM = theory of mind. [a]sum-score; [b]interference score; [c]maximum values are theoretically infinite; [d]proportion of correct switch trials; [e]proxy for information-processing capacity, standardized T-values.

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

Study 3

Table 2

Correlations of study variables

	EF	Conduct problems	Age	Information-processing capacity	Verbal ability	Gender
ToM	.41***	-.23***	.26***	.29***	.77***	-.14***
EF		-.04	.40***	.52***	.25***	.004
Conduct Problems			.01	-.05	-.15***	.23***
Age				-.05*	.33***	0
Information-processing capacity					.20***	0
Verbal ability						0
Gender						

Notes. EF = executive function, ToM = theory of mind, $N = 1{,}501$, gender was coded 1 = female, 2 = male.

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

Study 3

Figure Captions

Figure 1. Example stimulus display of the set-shifting task (adapted from Qu et al., 2013).

Figure 2. Structural equation model for hypotheses-testing analysis.

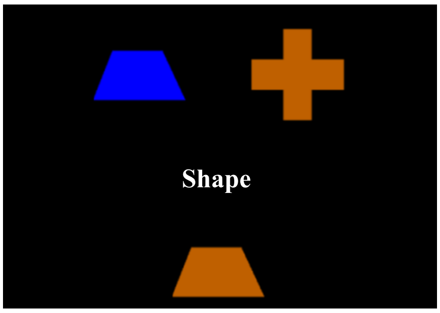
Notes. Standardized coefficients are displayed; dashed lines display non-significant paths; EF = executive function, ToM = theory of mind; all variables were controlled for age and gender, ToM was additionally controlled for verbal ability, and EF for information-processing capacity; the names of the first-order factors of the conduct problems factor are as followed: F1 = 'aggressive behavior', F2 = 'deviant behavior', F3 = 'devastation'.

N = 1,501; Model fit: $\chi^2(174) = 394.89$, $p < .001$; RMSEA = .03, 90% CI [.03, .03]; CFI = .91; SRMR = .04.

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

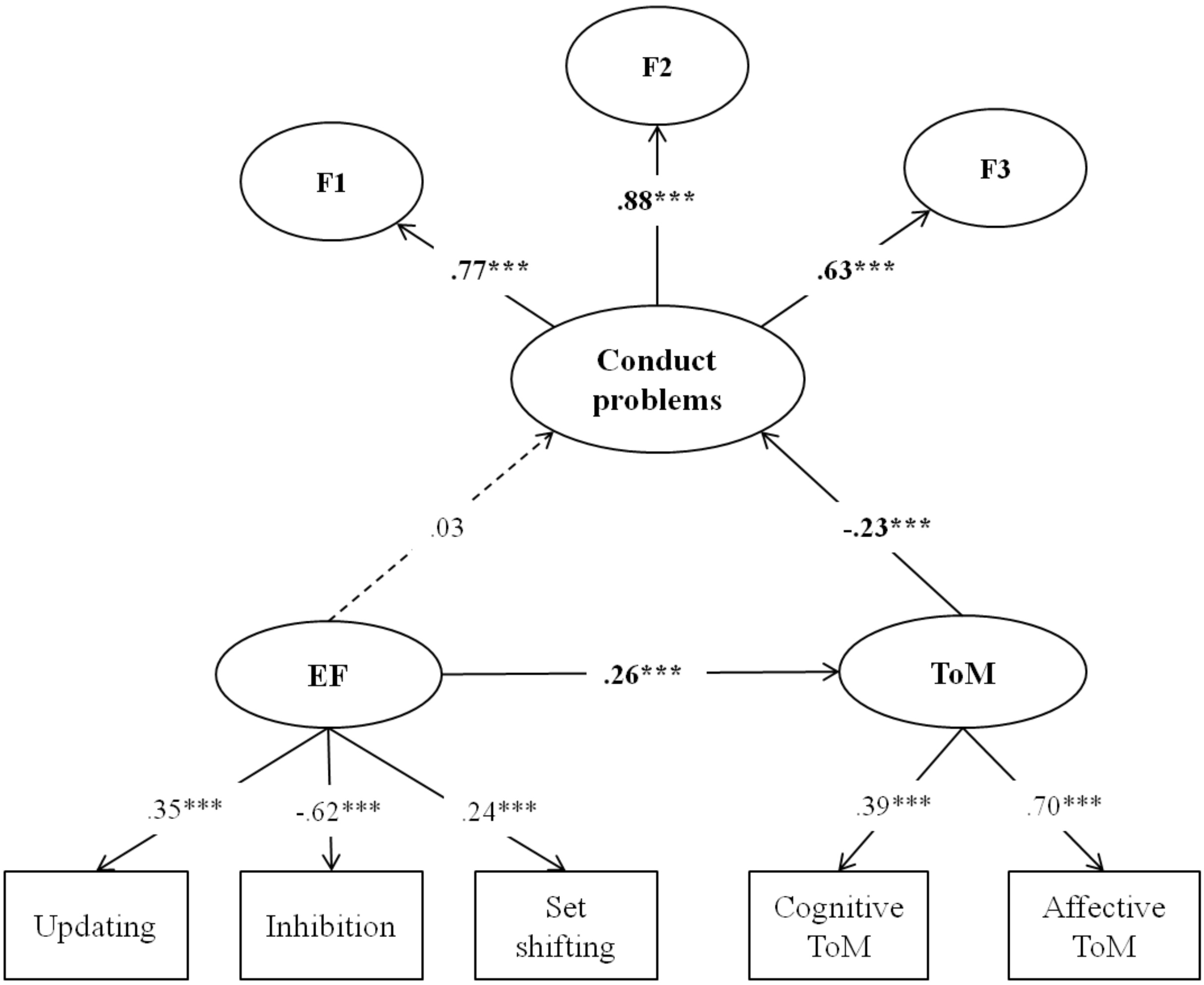
Study 3

Figure 1



Study 3

Figure 2



Study 3

5.1. Additional Figures for Study 3

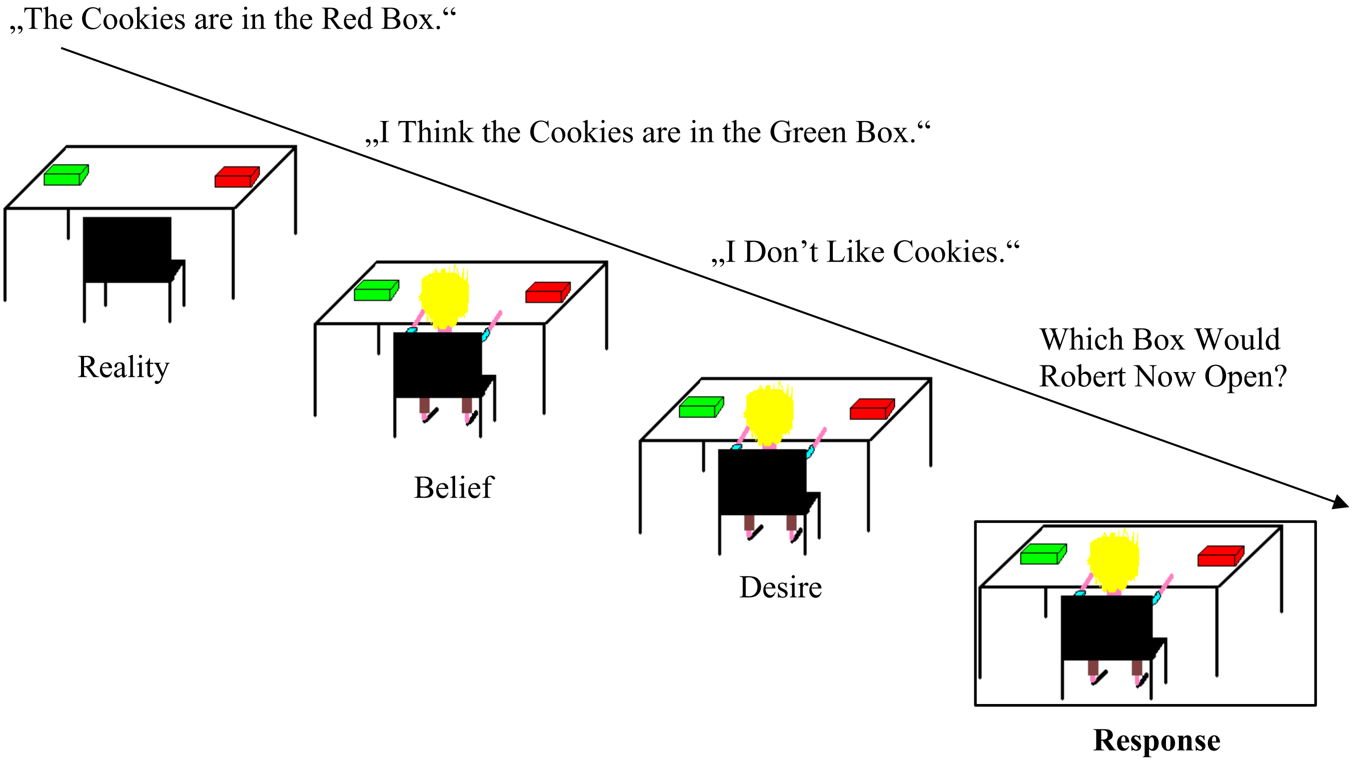
Figure Captions

Additional Figure 3. Example of chronological sequence of cognitive ToM task.

Additional Figure 4. Example picture from silent video clip and response choice screen of affective ToM task.

Study 3 – Additional Figures

Figure 3



Study 3 – Additional Figures

Figure 4

Silent Video Clip

1	2	3	4
restless	relieved	sullen	horrified

Choice Screen

6. Additional Analyses

Introduction

Moffitt (1993) proposed in the developmental taxonomy of antisocial behavior that the prevalence of antisocial behavior, which is a part of conduct disorder and problems, will increase during adolescence. This increase of antisocial behavior is explained mainly by the 'maturity gap' (Moffitt, 1993). Adolescents, who feel that they are biologically mature around the ages of 13-15 years, are not socially treated as "mature" adults until the age of 18 and for some aspects not until the age of 21 (see 2.1.2). Furthermore, in adolescence the influence of deviant peer group members increases, because they have found ways to cope with the discomfort of the maturity gap, for instance by delinquent behavior (e.g., Elliot & Menard, 1996). These deviant peers are assumed to show antisocial behavior from an early age on, and in adolescence they serve as role models for other adolescents. Previous evidence, for instance of the Dunedin longitudinal study (e.g., Moffitt et al., 2001), confirmed the proposition of increasing prevalence rates of antisocial behavior and conduct problems in adolescence.

Consequently, we wanted to study the developmental course of conduct problems in our longitudinal study, for which we used a latent growth curve model. In a latent growth curve model, developments can be modeled as growth between different measurement-points (e.g., Geiser, 2011). In a latent growth curve model, the latent intercept factor represents the true scores without measurement errors at the baseline level at T1, and the latent slope factor represents the difference of the true score variables over time. Additionally, the shape of the growth can be examined. Considering the theoretical considerations by Moffitt (e.g., 1993) and empirical evidence (e.g., Moffitt et al., 2001), we expected that for most of the children the conduct problems scores would be at a lower level at the beginning of our study at T1 (mean age of 8 years) as compared to later measurement points. Consequently, due to the approach to adolescence with increasing age of our children (mean age of 11 years at T3), we expected an increase over the course of our study, from T1 via T2 to T3. Therefore, we expected that the mean of the intercept of conduct problems in our sample at the beginning of our study, represented by the mean of the intercept factor, would be small (Ia) with low, but significant interindividual variability due to different developmental trajectories that could be present in our sample (Ib). Further, we expected a positive growth in conduct problem scores, that is, an increase in conduct problems, (IIa) with small interindividual variability between the children in our sample (IIb).

Theoretically, emotions and emotional regulation processes are supposed to mediate between social-cognitive processes and the social knowledge stored in the database, impacting on the behavior in a social situation (see section 2.1.7) (e.g., Lemerise & Arsenio, 2000). Inhibitory and emotional control are both processes involved in emotional regulation. Inhibitory control is conceptualized as a subdomain of effortful control, which is defined as the ability to inhibit a dominant response in order to perform a subdominant response, to detect errors, and to engage in planning (Rothbart & Rueda, 2005). Effortful control and EF have been determined as both relating to cognitive abilities that mainly overlap (Bridgett et al., 2013). Further, effortful control is proposed to predict emotional regulation in a social situation (Eisenberg & Spinrad, 2004). Consequently, inhibitory control, as a subdomain of effortful control, can also be related to another conceptualization of EF, namely hot EF. Zelazo and colleagues stated that EF processes and the neural systems that support EF vary as a function of motivational significance (Zelazo et al., 2010). 'Hot' EF are proposed to be top-down processes that operate in motivationally and emotionally significant situations, in contrast to 'cool' EF processes, which are supposed to operate in abstract, decontextualized situations that lack a significant affective or motivational component (e.g., Zelazo & Carlson, 2012). Hot and affective aspects of EF are usually associated with orbitofrontal cortex and other medial regions, and cool aspects are associated with lateral prefrontal cortex (e.g., Zelazo & Müller, 2011). Inhibitory control and emotional control also fall in this definition of hot EF, because they both consider aspects that include an affective or motivational component. However, the questionnaires that were used in our study do not represent typical ‘hot’ EF procedures.

Nevertheless, previous research determined that children with deficits in emotional control processes are prone to negative emotions and difficulties in regulating such emotions, and to deficits in inhibiting their behavior when emotionally aroused (e.g., for a review see Frick & Morris, 2004). These difficulties in emotion regulation are assumed to impair the development and use of social cognition, which is also assumed by the adapted SIP model (Lemerise & Arsenio, 2000). Empirically, deficits in emotional control or effortful control have been related to conduct problems by a large body of research in early and middle childhood (for a review, see Eisenberg, Spinrad, & Eggum, 2010). Consequently, deficits in emotional control and inhibitory control were proposed to predict higher levels of conduct problems in our study. So, we hypothesized negative paths from emotional control (III), and from inhibitory control (IV) to the intercept factor of conduct problems. Because of the before described theoretical propositions considering the relation between deficits in cool EF and higher levels of conduct problems, for instance the frontal-lobe hypothesis (e.g., Sèguin, 2009), also a negative path from cool EF to the intercept factor of conduct problems was proposed (V). The indicators for the cool EF factor were the three confirmed facets of EF, working memory updating, inhibition, and set-shifting, which were determined by Miyake et al. (2000) (see 2.3.1). Deficits in planning were proposed as an additional predictor of conduct problems, because of the proposition in the problem-solving account that deficits in planning would predict higher levels of aggressive behavior (e.g., Sèguin & Zelazo, 2005). So, a negative path was also proposed from the planning and organizing factor to the intercept factor of conduct problems (VI). Considering the prediction of the slope, this was examined as an open research question, because very little is known about the prediction of the trajectory of conduct problems in this age-period in a community sample. Considering early childhood, one study found increasing levels of conduct problems during kindergarten to first grade in an at-risk sample (Snyder, Schrepferman, McEachern, & Suarez, 2010). Theoretically, Moffitt (1993) proposed in her developmental taxonomy that children who have deficits in their EF ability would display higher levels of antisocial behavior, in particular from an early age-one (childhood-onset). However, later research showed that also adolescents whose antisocial behavior only starts with adolescence (adolescence-limited and -onset) display deficits in EF (Fairchild et al., 2013). Consequently, we assumed that deficits in all EFs, including cool EF, planning, and inhibitory, and emotional control, would lead to an increase in conduct problems over time, because they can represent a risk factor for conduct problems that start early, and also for conduct problems that only start with adolescence (VII).

Method

Participants

At the first measurement wave (T1) the sample included *N* = 1,647 children (52.09% females) between 6.23 and 11.33 years with a mean age of *M* = 8.36 years (*SD* = 0.95). At the second wave (T2), about 9 months after T1, *N* = 1,611 children (51.8% females) participated again (*M* = 9.12, *SD* = 0.93, range 7.12-11.90). At the third and final wave (T3), about 24 months after T2, 1,501 children (51.7% girls) participated again (*M* = 11.07, *SD* = 0.92, range 9.12-13.76). All children who participated at T1 were included in our analyses. Missing data were handled by the Full Information Maximum Likelihood procedure (Enders & Bandalos, 2001), and the actual number of observations for each study variable is noted in Table 1.

The study was part of the same larger research project on interpersonal developmental risk factors in childhood and adolescence from a longitudinal perspective as the previous studies. The children were recruited from 33 public primary schools in the Federal State of Brandenburg, Germany. The procedure and instruments were approved by the Ethics

Committee of the University of Potsdam, Germany, as well as by the Ministry of Education, Youth, and Sport of the Federal State of Brandenburg.

Material and Procedure

The data collection was similar like in the before described Studies 1-3 and took place at the children's schools. Parents of children that participated in our study were asked to fill in a questionnaire for each child.

Conduct Problems. We used the 5 items of the "Conduct Problems" scale of the parent form of the Strengths and Difficulties Questionnaire (Goodman, 1997) in the German version (Klasen et al., 2000). The parents were asked to rate their child's behavior on a three-point Likert-scale (0 = "not true", 1 = "partially/sometimes true", 2 = "very true") considering the last six months.

Inhibitory control. We used 6 items of the scale "Inhibitory control" of the Temperament in Middle Childhood Questionnaire (TMCQ) by Simonds (2006), for instance "My child can stop himself, when someone tells him to stop". The parents were asked to rate their child's behavior on a 6-point-Likert scale (1 = "not true", 2 = "rather not true", 3 = "partially true", 4 = "rather true", 5 = "very true") regarding the last six months.

Emotional control. We used 10 items of the scale "Emotional Control" scale of the Behavior Rating Inventory of Executive Function (BRIEF) by Goia and colleagues (Gioia et al., 2000), for instance "Has temper tantrums on smallest occasions". The parents were asked to rate their child's problems in emotional control on a 6-point-Likert scale (1 = never, 2 = rarely, 3 = sometimes, 4 = often, 5 = always) regarding the last six months. The items were recoded for the analysis, so that higher scores reflect higher emotional control.

Cool EF. The working memory updating, inhibition, and set-shifting tasks, as well as the planning scale were described in detail in the methods section of Study 2.

Additional Analyses

Information processing capacity. We used the Digit-Symbol-Test of the German version of the Wechsler Intelligence Scale for Children (Petermann & Petermann, 2007). For a detailed description see Study 2.

Statistical Analyses

We computed latent growth curve analyses over the three time-points using M*plus* (Version 7.4; Muthén & Muthén, 2012). First, a baseline latent growth curve model for conduct problems was determined. Because of the floor effects in our conduct problems measure, the conduct problems items were defined as censored. By doing this, during the modeling process it is automatically accounted for the restricted variance in the conduct problems scores. This is a particularly suitable way of modeling when floor or ceiling effects are encountered (Muthen & Muthen, 2012). For this way of modeling, a numerical integration algorithm is used, and as a consequence no normal fit indices are obtained, but a Loglikelihood, the Akaike (AIC), and the Bayesian information criterion (BIC). In all models, we used the robust Maximum Likelihood estimator (MLR) to account for non-normal distribution of our data, and all available data-points were included, but not all parents returned their questionnaires, so the sample for this analysis included $N = 1467$ children. In the model, in which the intercept and slope were predicted, information processing capacity, age and gender were covariates for all constructs. Effect sizes for *t*-tests are reported as Cohen's *d,* which were interpreted as small from .01 to .30, as moderate from .30 to .50, and large effects were > .50 (Cohen, 1988). For our SEM, we report standardized regression coefficients that can be interpreted in a similar way as Cohen's *d*., (Kline, 2011).

Results

Descriptive Results, Gender Differences, and Correlations

The descriptive statistics are presented in Table 1 as well as the results of statistical tests for gender differences. In the following only significant results are reported. Considering conduct problems, parent-rated scores were higher for boys as compared to girls at the first and third wave with a small effect size ($d \geq 0.15$). For behaviorally measured facets of cool EF, girls showed better set-shifting and inhibition abilities, but boys exhibited higher updating ability, all with a small effect size ($d \geq 0.03$). Girls showed higher parent-rated inhibitory control than boys with a small effect ($d = 0.26$), and in emotional control no difference emerged. Further, girls also exhibited higher information-processing capacity than boys with a medium effect ($d = 0.36$).

Considering the correlations, inhibitory control correlated positively in a moderate range with emotional control and planning, and negatively with conduct problems at all time-points. It further correlated with the behavioral EF measures, namely, positively with set-shifting and planning, and negatively with inhibition and gender, all in a small range. Similarly, emotional control correlated negatively with conduct problems at all three time-points in a moderate range, and positively with set-shifting and planning, as well as negatively with inhibition, all in a small range. Set-shifting and planning correlated positively with each other, and negatively with inhibition in a small to medium range. The negative correlations with inhibition stem from its measurement method, in which higher values display worse inhibition ability. Information-processing capacity correlated positively with set-shifting, updating, and emotional control in a small range, and in a medium range with planning, as well as negatively with inhibition and conduct problems at T1 in a small range. Age correlated positively with set-shifting and planning, and negatively with inhibition in a small to medium range and showed a small correlation with information-processing capacity. Gender displayed small and negative correlations with inhibitory control, set-shifting, planning and information-processing capacity, and positive with inhibition and conduct problems at T1.

Additional Analyses

Latent Growth Curve Model for Conduct Problems

In a latent growth curve model, the latent intercept factor represents the true scores without measurement errors at the baseline level at T1, and the latent slope factor represents the difference of the true score variables over time, here from T1 to T2 to T3. The slope factor can represent linear or, for instance quadratic growth. For conduct problems linear change was determined as adequate in the modeling process, because when quadratic growth was included in the model, this did not result in a solution, and linear change resulted in a good solution (see Figure 1).

The non-standardized regression coefficients for the intercept are defined as having a value of 1 at each time-point, so the individual values of each child in the intercept factor can take the value of the individual scores. The non-standardized regression coefficients for the linear slope factor represent the distance between the measurement points in months, because the T2 followed 9 months after T1, and T3 was conducted 36 months after T1.

The mean of the intercept factor displays the mean true scores of conduct problems in our study. The first hypothesis (Ia) was confirmed by a small mean of the intercept factor of conduct problems in our sample at T1, $\mu_i = 0.20$, $p < .001$. The mean of the true scores of conduct problems is very low in our sample, but significantly different from 0. The variance of the intercept is a measure for interindividual differences in the true scores of conduct problems. The second part of the hypothesis (Ib) considering the variance of the intercept factor was also confirmed in our sample at T1 by a small variance, $\sigma_{i2} = 0.12$, $p < .001$. The significance of the variance in our sample displays small, but significantly interindividual variability in the true scores of conduct problems at T1.

The mean of the latent slope factor represents whether growth across time-points occurred, and the variance displays whether interindividual differences in this growth occurred across time-points. The hypothesis about growth in conduct problems over the course of the study was not confirmed (IIa), because the mean of the slope factor in our sample has a negative value, $\mu_s = -0.09$, $p < .05$, which means that a small, but significant linear decrease in conduct problems occurred. This is illustrated by the estimated and observed sample means across all three time-points (see Figure 2). The second part of the hypothesis (IIb) was also not confirmed, because the variance of the slope factor in our sample is not significant, $\sigma_{s2} = 0.01$, $p = .75$, so no interindividual differences in the decrease across time occurred. Finally, we explored the covariance of the intercept and slope factor, but we had no directed hypothesis about it. The covariance was very small and not significant, *Cov(I, s)* = -.02, $p = .41$. Therefore, the decrease or change in the conduct problem scores of the children was not related to their baseline value at T1.

Prediction of Intercept and Slope of Conduct Problems

In a latent growth curve model, it is possible to predict the latent intercept and latent slope factors by other external variables. Different external variables were considered as predictors for the true scores of the children in conduct problems at the baseline level, that is, the latent intercept factor. We tested our hypotheses in a latent growth curve model (see Figure 4) that showed an acceptable fit (Loglikelihood = -52702.44; AIC = 105672.88; BIC = 106397.38), and in which we controlled all variables for information-processing capacity, gender, and age. In this model, planning could not be defined as an additional indicator for the cool EF-factor, because the model specification did not allow this. Therefore, it was defined as an additional latent factor (see Figure 3).

Our third hypothesis (III) was confirmed by a significant negative path from emotional control to the intercept factor of conduct problems, $\beta = -.38$, $p < .001$, and also the fourth (IV) hypothesis was confirmed by a significant negative path from inhibitory control to the intercept factor, $\beta = -.46$, $p < .001$, which were both of moderate size. Thus, the lower the child's emotional and inhibitory control was, the higher was the conduct problems score. Regarding the fifth hypothesis (V), the path from cool EF to the intercept factor was not significant, β = .06, p = .45. Thus, lower values of cool EF were not related to higher conduct problems scores. Finally, the sixth hypothesis (VI) was confirmed by a small, negative path from the Planning and Organizing factor to the intercept factor, β = -.18, $p < .001$.

Additionally, we examined the prediction of the slope. However, the slope was not predicted by any of the variables in our model, so the hypothesis (VII) that deficits in all EF would predict an increase in conduct problems was not confirmed.

Discussion

Altogether, a very low baseline level in conduct problems was determined, in which small interindividual differences were present in our sample at T1. This means that the majority of children in our study showed very low levels of conduct problems, and because of the small interindividual variability, some children might have had a little higher or lower levels of conduct problems. Further, a small linear decrease occurred over time, with no interindividual variability within the decrease, meaning that across all children the conduct problem scores in our study showed a small, but significant decrease. In accordance with the proposition of the developmental taxonomy, we expected increasing levels of conduct problems with the beginning of adolescence, because of the higher prevalence of adolescence-limited conduct problems (e.g., Moffitt, 1993). One reason, why we encountered decreasing instead of increasing levels could be that the last time-point was only at the transition to early adolescence, and with a mean age of 11 years, for some children adolescence might even not have started yet. Moffitt et al. (2001) determined that the adolescence-limited conduct problems were associated to the timing of puberty of each individual, because the conduct problems are mainly a consequence of the maturity gap experienced by the adolescents. Possibly, the majority of our children and young adolescents did not experience the discomfort of the maturity gap, yet, and consequently, did not have a reason to show conduct problems. We could not test this proposition, because we had no indicator for the starting of puberty in our study. A second reason for the decrease instead of an increase could be the floor effects in the conduct problem scores of our sample. Because the level of conduct problems was so small in our study altogether, it is difficult to interpret any change within.

Considering the prediction of the baseline level of conduct problems, some of our hypotheses were confirmed. The hypotheses on negative paths from emotional and inhibitory control to the intercept factor of conduct problems were confirmed. This means that deficits in emotional and inhibitory control were related to higher levels in conduct problems. This is in line with the propositions by Frick and Morris (2004) that children who have problems regulating their emotions and behavior are prone to show more conduct problems than other children. This is explained by their tendency to experience negative emotions, and their difficulties to control themselves in emotionally arousing situations, which both hinders the application of socio-cognitive skills. The result is also in line with the adapted SIP model by Lemerise and Arsenio (2000), which proposes that emotion processes, including emotional regulation processes, mediate between social-cognitive processes, the behavior in the social situation and the social knowledge. Negative emotions, such as anger have been found to mediate between social-cognitive processes, for instance hostile attributional biases and reactive aggression in children (e.g., De Castro, Merk, Koops, Veerman, & Bosch, 2005; Dodge et al., 2015). Our results are also in line with previous evidence on the relation between deficits in effortful control and conduct problems in middle childhood (e.g., Eisenberg et al., 2010).

Our hypothesis regarding the relation between deficits in cool EF and the conduct problems scores of our children was not confirmed. The frontal-lobe hypothesis states that children who show aggressive behavior display deficits in EF (Séguin, 2009). However, Séguin also proposed that the relation between physical aggression and EF might be even stronger for hot EF processes. This is in contrast to meta-analyses that found a relation between conduct problems and deficits for a broad range of EF facets (e.g., Ogilvie, Stewart, Chan, & Shum, 2011), however, in the meta-analysis also tasks that rather assess hot EF than cool EF were included. Consequently, the stronger relation between deficits in hot EF and conduct problems could have confounded the weaker or non-existent relation between deficits in cool EF and conduct problems in the meta-analysis. And when examined separately, cool, and hot EF could show differential relations to conduct problems. Moreover, Zelazo et al. (2010) argued that cool and hot EF processes always work together. Consequently, a task that measures cool EF processes always requires a certain amount of hot EF processes. Particularly in children, it is difficult to separate between cool and hot EF processes, because a test situation for them may always include some motivational aspects (e.g., Welsh & Peterson, 2014). Furthermore, it has to be noted that our conceptualizations of hot EF were parent-rated questionnaires, similar to the items for conduct problems, which were also rated by children's parents. Additionally, our conceptualization is different from tasks that are usually used to measure hot EF, for instance delay of gratification tasks, so it is difficult to compare our results to studies that used behavioral methods to measure hot EF. Therefore, the missing relation between deficits in cool EF and conduct problems could also be a consequence of missing overlap in the source of the data, because the cool EF were measured behaviorally.

This could also be the reason, why we found a path between deficits in planning and conduct problems, because the planning abilities were also not behaviorally measured, but rated by children's teachers. However, the result is also in line with the problem-solving account (Séguin & Zelazo, 2005), in which deficits in planning play an important role, because they hinder children's abilities to solve social situations adequately.

Altogether, we cannot determine, whether the result that only deficits in emotional and inhibitory control, which are related to hot EF processes, and planning ability were linked to higher conduct problems, but not cool EF, are a consequence of the difference between cool and hot EF or of the overlap in the source of the data. However, we found deficits in emotional and inhibitory control and planning abilities to be related to higher conduct problem scores in our children at T1, but their deficits did not predict the further development of their conduct problems across the course of our study. Maybe, additional, for example environmental variables like parenting practices influenced the developmental course of conduct problems (Eisenberg, Taylor, Widaman, & Spinrad, 2015). However, this cannot be determined with the data of our study. Another reason could be that the level of conduct problems was so low that the floor effects restricted the variance in our analyses, so that we did not find a prediction of the development of conduct problems despite controlling for the floor effects in our analyses.

In sum, future studies should disentangle the effects of deficits in cool, including planning, and hot EF in the relation to conduct problems by using multi-informant and multi-method measures. However, deficits in emotional and inhibitory control processes, as well as in planning seem to play an important role in conduct problems, and interventions preventing conduct problems should foster these abilities, which is possible by a variety of different and effective programs (Diamond & Lee, 2011).

Additional Analyses

References

Bridgett, D. J., Oddi, K. B., Laake, L. M., Murdock, K. W., & Bachmann, M. N. (2013). Integrating and differentiating aspects of self-regulation: Effortful control, executive functioning, and links to negative affectivity. *Emotion*, *13*, 47–63. http://doi.org/10.1037/a0029536

de Castro, B., Merk, W., Koops, W., Veerman, J. W., & Bosch, J. D. (2005). Emotions in social information processing and their relations with reactive and proactive aggression in referred aggressive boys. *Journal of Clinical Child and Adolescent Psychology*, *34*, 105–116. http://doi.org/10.1207/s15374424jccp3401

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, *333*(6045), 959–964. http://doi.org/10.1126/science.1204529

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

Eisenberg, N., & Spinrad, T. L. (2004). Emotion-related regulation: Sharpening the definition. *Child Development*, *75*, 334–339.

Eisenberg, N., Taylor, Z. E., Widaman, K. F., & Spinrad, T. L. (2015). Externalizing symptoms, effortful control, and intrusive parenting: A test of bidirectional longitudinal relations during early childhood. *Development and Psychopathology*, *27*(4pt1), 953–968. http://doi.org/10.1017/S0954579415000620

Elliot, D. S., & Menard, S. (1996). Delinquent friends and delinquent behavior: Temporal and developmental patterns. In J. D. Hawkings (Ed.), *Cambridge criminology series. Delinquency and crime: Current Theories.* (pp. 28–67). New York, NY: Cambridge

University Press.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*, 430–457. http://doi.org/10.1207/s15328007sem0803_5

Fairchild, G., Van Goozen, S. H. M., Calder, A. J., & Goodyer, I. M. (2013). Research review: Evaluating and reformulating the developmental taxonomic theory of antisocial behaviour. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *54*, 924–940. http://doi.org/10.1111/jcpp.12102

Frick, P. J., & Morris, A. S. (2004). Temperament and Developmental Pathways to Conduct Problems. *Journal of Clinical Child and Adolescent Psychology*, *33*, 21–31. http://doi.org/10.1207/S15374424JCCP3301

Geiser, G. (2011). *Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung* (2.). Wiesbaden: VS Verlag für Sozialwissenschaften.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior Rating of Executive Function.* Lutz, FL: Psychological Assessment Resources.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *38*, 581–586.

Klasen, H., Woerner, W., Wolke, D., Meyer, R., Overmeyer, S., & Kaschnitz, W. Rothenberger, A. Goodman, R. (2000). Comparing the German versions of the Strengths and Difficulties Questionnaire (SDQ-Deu) and the Child Behavior Checklist. *European Child and Adolescent Psychiatry*, *9*, 271–276.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.

Lemerise, E. A., & Arsenio, W. F. (2000). An integrated model of emotion processes and cognition in social information processing. *Child Development*, *71*, 107–118.

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

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide. Statistical analysis with latent variables* (7th ed.). Los Angeles, CA: Muthén & Muthén.

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*, 1063–1107. http://doi.org/10.1111/j.1745-9125.2011.00252.x

Petermann, F., & Petermann, U. (Eds. . (2007). *Hamburg-Wechsler-Intelligenztest für Kinder - IV (HAWIK-IV) [Wechsler intelligence scale for children IV].* Göttingen: Hogrefe.

Rothbart, M. K., & Rueda, M. R. (2005). The development of effortful control. In U. Mayr, E. Awh, & S. Keele (Eds.), *Developing individuality in the human brain: A tribute to Michael I. Posner* (pp. 167–188). Washington, D.C.: American Psychological Association.

Séguin, J. R. (2009). The frontal lobe and aggression. *European Journal of Developmental Psychology*, *6*, 100–119. http://doi.org/10.1080/17405620701669871

Simonds, J. (2006). *The role of reward sensitivity and response: Execution in childhood extraversion.* Oregon.

Snyder, J., Schrepferman, L., McEachern, A., & Suarez, M. (2010). Early Covert Conduct Problems: Phenomenology , Prevalence , Cross-Setting Diffusion , Growth and Consequences. *The Journal of Behavior Analysis of Offender and Victim Treatment and Prevention*, *2*(1), 4–19. http://doi.org/10.1037/h0100471

Welsh, M., & Peterson, E. (2014). Issues in the conceptualization and assessment of hot executive functions in childhood. *Journal of the International Neuropsychological Society*, *20*, 152–156. http://doi.org/10.1017/S1355617713001379

Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, *6*, 354–360. http://doi.org/10.1111/j.1750-8606.2012.00246.x

Zelazo, P. D., & Müller, U. (2011). Executive function in typical and atypical development. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (2nd ed., pp. 574–603). Chichester, UK: Whiley-Blackwell. http://doi.org/10.1002/9781444325485.ch22

Zelazo, P. D., Qu, L. I., & Kesek, A. C. (2010). Hot executive function: Emotion and the development of cognitive control. In S. D. Calkins & M. A. Bell (Eds.), *Child development at the intersection of emotion and cognition* (pp. 97–111). Washington, DC: American Psychological Association. http://doi.org/10.1037/12059-006

Additional Analyses

Table 1

Descriptive Statistics

	n (*n* female)	Range	Total *M* (*SD*)	Girls *M* (*SD*)	Boys *M* (*SD*)	Gender difference effect size
Emotional Control[a]	1338 (691)	1.5 - 5	3.79 (0.71)	3.79 (0.69)	3.80 (0.73)	$t(1336) = 0.10$ $d = 0.01$
Inhibitory Control[a]	1335 (773)	1.17 - 5	3.53 (0.67)	3.61 (0.64)	3.44 (0.69)	$t(1308.63) = 4.64$*** $d = 0.26$
Updating[b]	1641 (855)	0 - 13	6.18 (1.47)	6.16 (1.46)	6.20 (1.48)	$t(1639) = 0.60$*** $d = 0.03$
Inhibition[c]	1647 (860)	7.07 - 89.03[c]	24.93 (8.77)	24.26 (8.0)	25.68 (9.49)	$t(1543.54) = 3.26$*** $d = 0.16$
Set-Shifting[d]	1640 (855)	0 - 22	15.57 (4.68)	16.21 (4.68)	14.88 (4.58)	$t(1638) = 5.78$*** $d = 0.29$
Planning[a]	1419 (739)	1 - 5	3.70 (0.89)	3.89 (0.85)	3.50 (0.90)	$t(1417) = 8.27$*** $d = 0.44$
Conduct Problems T1[a]	1339 (692)	0 - 1.80	0.30 (0.31)	0.27 (0.29)	0.33 (0.33)	$t(1290.81) = 3.49$*** $d = 0.19$
Conduct Problems T2[a]	1193 (620)	0 - 1.80	0.27 (0.29)	0.26 (0.29)	0.29 (0.29)	$t(1191) = 1.56$ $d = 0.09$
Conduct Problems T3[a]	1069 (556)	0 - 1.60	0.27 (0.30)	0.25 (0.29)	0.30 (0.31)	$t(1067) = 2.50$* $d = 0.15$
Digit-symbol-test[e]	1647 (858)	27 - 80	51.20 (9.52)	52.82 (9.57)	49.44 (9.15)	$t(1645) = 7.32$*** $d = 0.36$

Notes. [a]mean-score; [b]sum score; [c]interference score, maximum values are theoretically infinite; [d]number of correct switch trials; [e]proxy for information-processing capacity, standardized T-values.

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

Additional Analyses

Table 2

Correlations of study variables

	T1 EC	T1 CP	T2 CP	T3 CP	T1 U	T1 S	T1 I	T1 P	T1 DST	Age	Gender
T1 IC	.40***	-.48***	-.42***	-.37***	.09*	.14***	-.12***	.29***	.05	.05	-.13***
T1 EC		-.54***	-.44***	-.38***	.09**	.06*	-.10***	.16***	.08**	-.02	0
T1 CP			.61***	.54***	-.09***	-.13***	.10***	-.32***	-.10***	.01	.10***
T2 CP				.56***	-.07*	-.14***	.11***	-.24***	-.04	-.02	.05
T3 CP						-.05*	.08*	-.26***	-.01	.01	.08*
T1 U						.35***	-.27***	.27***	.12***	.24***	.02
T1 S							-.33***	.30***	.16***	.28***	-.14***
T1 I								-.28***	-.26***	-.35***	.09***
T1 P									.29***	.02	-.21***
T1 DST										-.08**	-.18***
Age											.03

Notes. IC = Inhibitory Control, EC = Emotional Control, CP = Conduct Problems, U = Updating, S = Set-Shifting, I = Inhibition, P = Planning, DST = Digit Symbol Test, *N* = 1,501, gender was coded 1 = female, 2 = male.

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

Additional Analyses

Figure Captions

Figure 1. Latent growth curve model of conduct problems.

Notes. Non-standardized regression coefficients are displayed; dashed lines display non-significant paths; N = 1467; Model fit: Loglikelihood = -1970.56; AIC = 3957.13; BIC = 3999.45.

Figure 2. Estimated and observed sample means.

Figure 3. Latent Growth curve model with prediction of intercept and slope.

Notes. EC = Emotional Control, PO = Planning and Organizing, IC = Inhibitory Control. Standardized regression coefficients are displayed; dashed lines display non-significant paths; all variables were controlled for information-processing capacity, gender, and age; $N = 1647$; Model fit: Loglikelihood = -52702.44; AIC = 105672.88; BIC = 106397.38.

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

Additional Analyses

Figure 1

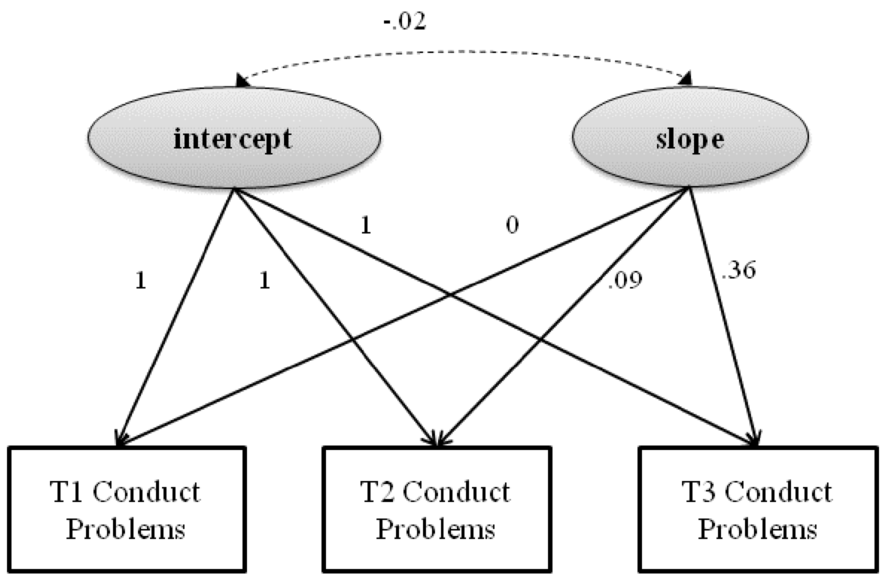


Figure 2

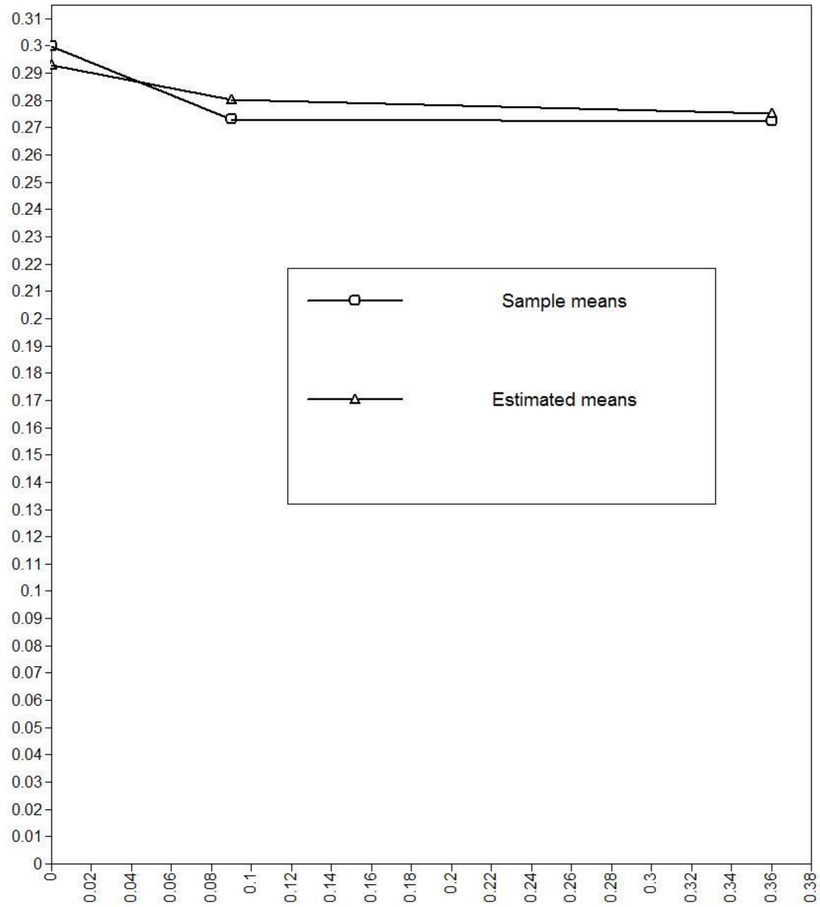
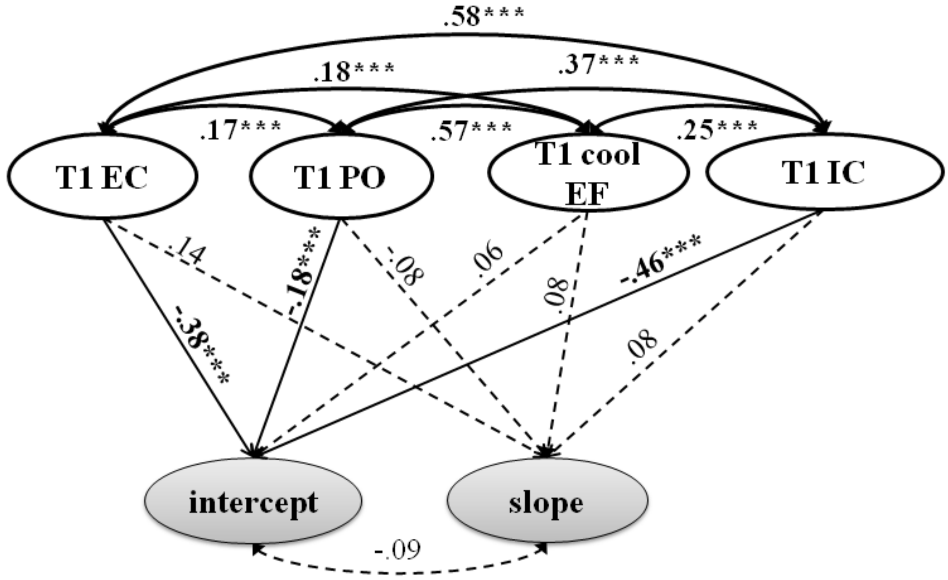

0.31
0.3
0.29
0.28
0.27
0.26
0.25
0.24
0.23
0.22
0.21
0.2
0.19
0.18
0.17
0.16
0.15
0.14
0.13
0.12
0.11
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16
0.18
0.2
0.22
0.24
0.26
0.28
0.3
0.32
0.34
0.36
0.38
Sample means
Estimated means

Figure 3


.58***
.18***
.37***
.17***
.57***
.25***
T1 EC
T1 PO
T1 cool EF
T1 IC
.14
-.18***
-.08
.06
.08
-.46***
-.38***
.08
intercept
slope
-.09

7. Discussion

7.1. Main discussion and research questions

This thesis investigated whether deficits in ToM and EF are risk factors for conduct problems from middle childhood through early adolescence. The theoretical background mainly centered around the SIP model by Crick and Dodge (1994), which proposes that children who show aggressive behavior have deficits in their processing of social cues, among other theoretical accounts. This biases children's social cognition in a way that leads them to behave aggressively towards others. Further, reciprocal processes were also proposed in this thesis, which means that reverse pathways from earlier aggressive behavior to later ToM were also expected. This proposition is supported by the SIP model and newer transactional models, in particular the ontogenic processes model of externalizing psychopathology by Beauchaine et al. (2015). To evaluate these expectations, a longitudinal perspective was adopted, which is also in line with the developmental taxonomy by Moffitt (1993). The developmental periods of middle childhood and early adolescence were determined to be particularly appropriate for this research, because of simultaneous changes in social, cognitive and emotional domains (e.g., Hartup, 1984; Steinberg & Morris, 2001).

Three main research questions were examined in the thesis:

I) Are deficits in ToM a risk factor for conduct problems, including different forms of aggressive behavior, from middle childhood through early adolescence?

II) Are deficits in EF a risk factor for conduct problems, including different forms and functions of aggressive behavior, from middle childhood through early adolescence?

III) Are deficits in EF and ToM risk factors for conduct problems from middle childhood through early adolescence when they are studied in tandem?

An additional aim of the research was to assess the moderating role of age and gender differences in the examined relations.

To investigate these questions, three studies were conducted. All studies were based on the same sample of children, which was part of a larger longitudinal research project on intrapersonal risk factors for child and adolescent psychopathology at the University of Potsdam, although the three studies used different measurement points (T1, T2, and T3) and different variables of the community sample.

Altogether, the assumptions of the SIP model could be confirmed to a great extent. It was found that deficits in ToM were indeed related to a greater number of conduct problems, including aggressive behavior. The proposed reverse pathway from aggressive behavior to later deficits in ToM was only confirmed for relational aggression. This will be explained in detail in section 7.2.

The role of deficits in EF as a risk factor for conduct problems was confirmed for relational, physical, and reactive aggression, as well as conduct problems. A detailed discussion of these links follows in section 7.3. Then, the results of the study of deficits in ToM and EF in tandem as risk factors for conduct problems will be discussed in section 7.4. Afterwards, limitations of the discussed studies and the present thesis will be outlined (7.5.), and possible implications for research and practice will be debated (7.6.). Finally, a conclusion for the thesis will be drawn (paragraph 8.).

7.2. Deficits in ToM as a risk factor for conduct problems

The first research question for the present thesis concerned deficits in ToM as a risk factor for conduct problems, including different forms of aggressive behavior. This relation was investigated in Studies 1 and 3 (paths a and b, Figure 2, p. 56). Theoretically, deficits in ToM are proposed to be involved in deviant processing of social cues, for instance in the attribution of intent, which leads to aggressive behavior (Crick & Dodge, 1994). Further, deficits in ToM are also assumed to hinder an important mechanism inhibiting violent and antisocial behavior (e.g., Sharp, 2008). Previous research has yielded mixed evidence: a large body of studies has yielded a negative relation (e.g., Capage & Watson, 2001), whereas some studies have found a positive relation, in particular in the context of relational forms of aggression (e.g., Renouf, 2010). In addition, two studies yielded results that are difficult to interpret, although they found no direct correlation between conduct problems and deficits in simple ToM tasks (Happé & Frith, 1996; Hughes et al., 2000). One reason for the equivocal evidence regarding this relation may be that the studies investigated either physical or relational aggression or did not differentiate between these two forms of aggressive behavior. Since aggressive behavior is a major part of conduct problems, differentiation between forms of aggression, namely, physical and relational, is deemed important. Consequently, this was investigated in Study 1.

Previous research on the relation between physical aggression and ToM has mainly found a negative association (e.g., in 3- to 7-year-olds: Capage & Watson, 2001). However, until now this had mainly been investigated in early childhood (e.g., Hughes et al., 1998), and previous research was mostly cross-sectional. Studies examining the association between deficits in ToM and conduct problems in middle childhood and early adolescence did not differentiate between physical and relational aggression (in 7- to 10-year-olds: Anastassiou-Hadjicharalambous & Warden, 2008; in 5- to 9-year-olds: Fahie & Symons, 2003; in 9- to 13-year-olds: Sharp, 2008), but they all yielded a negative association. Therefore, we hypothesized that – based on the SIP theory by Crick and Dodge (1994) – deficits in earlier ToM would predict later, higher scores in physical aggression. As a result, our prediction in Study 1 was confirmed by a negative path from ToM at T1 to physical aggression at T2, and by a negative path from ToM at T2 and physical aggression at T3. These results are consistent with previous cross-sectional evidence (e.g., Anastassiou-Hadjicharalambous & Warden, 2008). Further, as this was a longitudinal study, we controlled for earlier levels of aggressive behavior, and also partialled out stable between-person differences in aggressive behavior. In addition, we controlled ToM for verbal ability, because the two are interrelated throughout development (Milligan et al., 2007). Consequently, deficits in ToM were determined as a risk factor for physical aggression over a period of three years on an intrapersonal level from middle childhood to the transition to early adolescence. This can be explained in terms of the SIP model, because as a consequence of their inability to infer the cognitive and affective mental states of others correctly, children in our study with earlier deficits in ToM might have displayed deficits and deviances in their cognitive SIP. For instance, they could have shown problems in the attribution of intent, which is part of encoding and interpretation, Steps 1 and 2, of the SIP model. Their failure to understand the motives of others, such as their peers, leads them to behave physically aggressively towards them, because, for instance, they tend to interpret a hostile intent to others' actions. Therefore, it might be that deficits in ToM and deficits in SIP are also related over the course of middle childhood through early adolescence longitudinally, as Choe et al. (2013) have shown in their study for this link from early childhood to school age. This longitudinal developmental link might foster the development of physically aggressive behavior over time. However, this must be proven in future longitudinal studies.

Another possible interpretation is in accordance with the explanations by Anastassiou-Hadjicharalambous and Warden (2008) and Sharp (2008), who also found a negative relation between ToM and physical aggression or conduct problems, respectively. They hypothesized that ToM is a prerequisite to empathic responding, which is proposed to be involved in an important mechanism inhibiting violent and antisocial behavior. As deficits in ToM and empathy are highly associated (e.g., Shamay-Tsoory et al., 2003) – both involve activation of the amygdala, for example – a consequence of that could be violent behavior against others. Further, the ability to infer another's perspective is important for moral conduct (e.g., Hoffman, 1987), and deficits in the ability to infer and understand another's cognitive and affective perspective can foster aggressive behavior towards others (Anastassiou-Hadjicharalambous & Warden, 2008; Hare, 1970).

Regarding relational aggression, previous evidence is limited and mixed. One longitudinal study found a positive link between ToM and relational aggression in preschool-aged children over a period of one year (Renouf et al., 2010). In middle childhood, a cross-sectional study determined that the association between relational aggression and ToM was positive for children aged 7.6 years or younger, and negative for children aged 9.6 years or older, and not significant in between (Gomez-Garibello & Talwar, 2015). Further, in early adolescence, a negative association was found between relational aggression and ToM (Kokkinos et al., 2016). Thus, we examined the association between ToM and relational aggression as an open research question. As a result, in our study we determined a negative path from earlier deficits in ToM to later relational aggression; this was true for the path from ToM at T1 to relational aggression at T2, as well as for the path from ToM at T2 to relational aggression at T3. These results were obtained under the control of earlier levels of relational aggression, and after partialling out between-person differences in aggressive behavior. ToM was also controlled for verbal ability. So, we determined deficits in ToM as a risk factor for later relationally aggressive behavior on an intrapersonal level over a period of three years. This result is also in line with the propositions of the SIP model. It confirms that children with deficits in ToM may show deviant cognitive processing in SIP, which leads them to show more relationally aggressive behaviors towards their peers one and two years later. Our findings can be interpreted in such a way that children who had earlier problems inferring other persons' cognitive and affective mental states later threatened other children through relational means, for instance, by trying to exclude certain peers from peer group activities. One reason for this longitudinal link could be that children with deficits in inferring others' cognitive and affective mental states tend to misinterpret the motives and intentions of others, such as interpreting a hostile intent to them. They may, for example, be unable to infer the peer's belief about reality, and therefore might not be able to understand that different beliefs about reality exist. Consequently, they could try to protect themselves against the perceived hostile intentions of others by excluding them from their peer group. Another possible reason could be that children with deficits in their ToM ability might not be able to accurately predict a certain expected outcome after their response to the situation at hand. Consequently, they might not adjust their response decision adaptively, as in Step 5 of the SIP model, in accordance with peer behavior in a given situation, as has been shown by Grueneisen et al. (2015). As mentioned before, a developmental link between deficits in ToM and deficits in SIP could be responsible for the development of relationally aggressive behavior. However, this is a topic for future research.

Our result is in accordance with the study by Kokkinos et al. (2016) that also determined a negative link between relational aggression and ToM. It is partially in line with the result of Gomez-Garibello and Talwar's (2015) study, because the children in our study were around the age of 8 years at T1, 9 years at T2, and 11 years at T3. In their study, a negative link was only determined for children aged 9.6 years and older. Nevertheless, both studies were cross-sectional, and consequently cannot be compared perfectly to our longitudinal study. However, our result does not support the proposition by Sutton et al. (1999a) that children with superior ToM use their proficient ability to infer others' mental states in order to manipulate them. Although this proposition was – until now – mainly confirmed for bullying behavior, particularly ring-leader bullying (Sutton et al., 1999b), and one other study. In the other study – apart from Sutton et al. (1999b) – that also found a positive relation between indirect aggression, which is very similar to relational aggression, and ToM, prosocial behavior moderated the relation (Renouf et al., 2010). This means that only children with average or low levels of prosocial behavior used their proficient ToM ability to manipulate other children through indirect forms of aggression, but not children who showed high levels of prosocial behavior. In our study, we did not investigate a moderation by prosocial behavior, and it could be that this is the reason for the contrasting results. Another reason could be that the items in Renouf et al.'s study seem to be more similar to indirect forms of bullying, and as such the proposition by Sutton et al. might be more appropriate for them. Another difference between our study and Renouf et al.'s study is that they did not control for earlier levels of aggressive behavior despite examining the relation longitudinally and did not partial out between-person differences in aggression.

Further, in our study, we investigated reciprocal relations between ToM and aggressive behavior. This means that we also examined the reverse path from aggressive behavior to ToM, which was already included theoretically in the social information-processing model (Crick & Dodge, 1994), and is also proposed in recent transactional models, like the ontogenic processes model of externalizing psychopathology (Beauchaine et al., 2015). In both theoretical approaches, it is assumed that a cognitive deficit may elicit certain responses from others, particularly peers, mediated through the child's aggressive behavior towards them, which then feeds back into the child's development and behavior. In line with the SIP model, this means that past social experiences that are stored in long-term memory are retrieved at every step of SIP from the 'dataset.' Because ToM is involved in at least some of these steps, it might also be influenced by memories of past aggressive behavior. In light of these theoretical propositions, we hypothesized that higher levels of earlier aggressive behavior would lead to later deficits in children's ToM. As a result, we found a negative path from relational aggression at T1 to ToM at T2. Hence, our assumption could be confirmed for relational aggression in our study, but not for physical aggression. We found reciprocal processes between deficits in ToM and relational aggression across time points from T1 relational aggression to T2 ToM to T3 relational aggression. This can be interpreted in such a way that the relational aggression feeds back into the children's ToM ability through a developmental link over time, and hinders further development of ToM, which would be appropriate at this age. This finding is also in line with the assumption by Hughes and Leekam (2004) that social experiences influence the development of ToM in either a positive or a negative way. Possibly, the children who showed higher levels of relational aggression at the beginning of our study already displayed deficits in their ToM before the beginning of our study, which led them to behave aggressively at T1. This characterizes the developmental transactions between a cognitive deficit, in our case a deficit in ToM ability, and the behavior of the children and their resulting cognitive and social development, which were proposed in the SIP model (Crick & Dodge, 1994) and the ontogenic processes model (Beauchaine et al., 2015). One previous study detected a developmental cascade between SIP and aggressive behavior in children (Lansford et al., 2010), though they did not investigate ToM ability in particular. It is possible that SIP and ToM ability might both be intertwined with aggressive behavior in these developmental cascades. However, this has to be proven in future research.

Our result is in line with the only prior study that investigated the reverse path from aggressive behavior to ToM and also found a negative path from earlier aggression to later ToM in preschool age (Song et al., 2016). However, they assessed child aggression in the form of problem behaviors and did not differentiate between physical and relational aggression, which is in accordance with the studied age range of early childhood, in which relational aggression is not as developed as physical aggression.

One reason that could be responsible for the missing path from physical aggression to ToM may be that different functions of aggression, namely, reactive and proactive, could lead to differing relations to ToM. However, research on these relations is rare, and a recent study on a sub-sample of the PIER study found no longitudinal paths from earlier reactive and proactive aggression to later affective or cognitive ToM (Austin, Bondü, & Elsner, 2017). Considering different functions of physical aggression, until now, only reactive physical aggression has been related to low socio-cognitive abilities (Crick & Dodge, 1996), not proactive physical aggression, which might instead show a positive relation due to its instrumental character. Therefore, these competing associations could have cancelled each other out, and thus eliminated the path from physical aggression to ToM. Additional reasons could be that we did not consider other socio-cognitive processes, like encoding, as a part of the SIP model, or moderating variables like prosocial behavior (Renouf et al., 2010).

Further, we also investigated whether the reciprocal paths between ToM and physical and relational aggression in our model were moderated by gender and age. Therefore, we computed additional multi-group analyses, which showed that our results remained equal for girls and boys, and three different age groups. This suggests that despite differences in frequencies in some of the study variables between females and males in our sample, for instance with boys showing more physically and relationally aggressive behavior, the underlying mechanisms between deficits in ToM and physically and relationally aggressive behavior are similar between the two sexes. The result that boys showed more physical aggression than girls is in line with previous research (e.g., Lansford et al., 2012); however, research on gender differences in relational aggression is mixed. Some studies have found that girls show more relational aggression than boys (e.g., Ostrov et al., 2013), whereas others found no or only a trivial gender difference (e.g., Card et al., 2008), and one study found more relationally aggressive behavior in boys (Henington et al., 1998). Therefore, the result that the boys in our study showed more relational aggression than the girls at two time points is in line with one earlier study (Henington et al., 1998). The age of the children in Henington's study (1998) might be similar to the mean age of about 9 years for our children at T2, however, they only reported that their children were in the second or third grade of elementary school.

Discussion

Physically and relationally aggressive behaviors are highly correlated in individuals (see Card et al., 2008), and as such the small gender difference in relational aggression could also be an indirect consequence of the large gender difference in physical aggression in our study. This is in line with the observation that the correlation between physical and relational aggression had to be freely estimated in our multi-group model, indicating a gender difference in this correlation with boys showing a higher correlation than girls. Considering the ToM scores, no significant gender difference was found, although girls displayed minimally higher scores than boys, which is in line with previous studies (e.g., Devine & Hughes, 2013). Girls' superiority on ToM tasks is usually explained by different play behaviors, with girls favoring activities that foster social interaction in contrast to boys, who typically prefer activities that foster mainly spatial abilities. Other explanations encompass biological differences between the two gender groups, with females, for instance, showing superiority in empathic abilities or abilities that are related to empathy, like ToM (e.g., Baron-Cohen, 2002). It could be a consequence of the ceiling effect in our ToM task that the gender difference was not larger or significant. Despite these mean level differences, the paths between ToM and both physical and relational aggression did not differ between girls and boys. This is an interesting result, because very little is known about whether mean-level differences between the two sexes really affect the underlying mechanisms and effects between deficits in ToM and conduct problems, including physical and relational aggression. Consequently, our results suggest that the mean-level differences in the involved variables are not decisive for differences between the sexes in the underlying effects of ToM and both forms of aggression.

Another question was whether the effects also remain similar from middle childhood through early adolescence despite reported differences in frequency levels across different age groups, such as an increase in conduct problems during adolescence (e.g., Moffitt et al., 2001). Other research has found that ToM ability also increases during middle childhood and adolescence (e.g., Apperly et al., 2011; Vetter et al., 2013). As a result, we found that the reciprocal paths between ToM and physical and relational aggression remained equal when comparing the three equally sized age groups of our study from a mean age of 8 to 11 years (6.23 to 7.83 years, 7.84 to 8.86 years, 8.87 to 11.3 years). However, we found a significant correlation between ToM and age, which indicates that ToM ability is still improving in this age range, and which is in line with previous research (e.g., Apperly et al., 2011). Our results suggest that the correlation of ToM with age is not decisive for differences in the underlying mechanisms between deficits in ToM and physical and relational aggression across different age groups.

In Study 3, the path from ToM to conduct problems was examined cross-sectionally. Due to ceiling effects in the ToM measure in our longitudinal study, the PIER study, the assessment method of ToM had to be changed at the third measurement point. Consequently, in this case, a longitudinal modeling which included ToM from the first to the third time point was not appropriate. Therefore, we only analyzed data from the third measurement point in Study 3, in which the children were in the transition from middle childhood to early adolescence. Evidence on the association between ToM and conduct problems in particular is mixed, and most evidence comes from studies with preschool-aged children (e.g., Hughes et al., 1998), while some evidence comes from studies in middle childhood (e.g., Sharp, 2008). In Hughes et al.'s (1998) study, hard-to-manage preschoolers performed worse on ToM tasks compared to children with low levels of conduct problems, however, their cognitive and affective ToM scores did not correlate with their conduct problems (Hughes et al., 2000). Likewise, Happé & Frith (1996) also found a mixed result. In their study, the children with higher levels of conduct problems mastered simple ToM tasks with no difference from normally developing children. However, the children with higher levels of conduct problems were rated to have deficits in the application of their ToM abilities in everyday life. Thus,

Discussion

Happé and Frith proposed (1996) that the children who showed higher levels of conduct problems might use their ToM abilities in antisocial ways, for instance, by cheating and lying. In a similar direction argued Sutton et al. (1999b), who proposed that children with superior ToM might use their ability to manipulate others; however, they investigated only bullying behavior, thus their proposal cannot be applied to the present study. Considering the developmental periods of middle childhood and early adolescence, two other studies investigated the link in this age range. Both found a negative link (Anastassiou-Hadjicharalambous & Warden, 2008; Sharp, 2008). Different reasons have been accounted for the mixed evidence: First, the ease of many ToM tasks that do not assess abilities relevant for everyday social interactions (Richell et al., 2003) could be responsible. Second, the failure to differentiate between the cognitive and affective facets of ToM (Anastassiou-Hadjicharalambous & Warden, 2008) or, third, the failure to control ToM for its relation to verbal abilities could be other reasons (Milligan et al., 2007). Consequently, we applied a cognitive (Apperly et al., 2011) and an affective ToM (Vetter et al., 2014) task that have been shown to detect individual and developmental differences in this age range. In addition, the Faces task (Vetter et al., 2014) that we used to assess affective ToM shows some similarities to another task, the "reading the emotions of the eyes task," which revealed deficits in ToM in children with higher levels of conduct problems (Sharp, 2008). And, last, we controlled the ToM scores for verbal ability.

Theoretically, we also based our hypotheses on the SIP theory (Crick & Dodge, 1994) to explain the relation between deficits in ToM and aggressive behavior in conduct problems, like in Study 1 described above. To explain the antisocial behavior of children with conduct problems, we argued in terms of the amygdala theory of violent behavior, which states that deficits in ToM are related to deficits in empathy, which hinder an important mechanism inhibiting antisocial and violent behavior (e.g., Blair et al., 2014; Sharp, 2008). We hypothesized in line with the SIP theory and the amygdala theory of violent behavior that deficits in ToM would be related to higher levels of conduct problems. According to the results, the hypothesized relation between deficits in ToM and more conduct problems was confirmed in Study 3. This is consistent with the propositions of the SIP model and the amygdala theory of violent behavior, because the children who had deficits in inferring the affective and cognitive mental states of others showed more conduct problems, such as threatening others. One explanation for this link could be that they had deficits in their ToM that are closely related to deficits in empathy and that these hindered the violence-inhibiting mechanism. Further, their inability to take another's cognitive and affective perspective made it difficult for them to behave in morally appropriate ways. Therefore, they behaved in a violent and antisocial way towards others. Our finding is also in line with two of the earlier studies in the same or a similar age period (Anastassiou-Hadjicharalambous & Warden, 2008; Sharp, 2008). The study by Sharp (2008) also investigated a community sample with preadolescents and young adolescents between 9 and 13 years, which is the same as our age range in this study. However, they only used an affective ToM measure, the "reading the emotions of the eyes task" for children, which is only partially similar to ours. We also asked our participants to identify an emotion or higher affective state in our affective ToM task, but showed them short, silent video clips that included the face and torso of each actor. In our study, we additionally used a cognitive task that was also integrated in the latent ToM factor; however, the factor loading of the affective ToM task was substantially higher, and consequently had more weight in the integrated factor. Altogether, both studies identified a deficit in ToM in children with higher levels of conduct problems in a community sample aged between 9 and 13 years. This congruence points to a possible general deficit in ToM, particularly in the affective facet, in children with higher levels of conduct problems.

Discussion

Considering the results by Anastassiou-Hadjicharalambous and Warden (2008), our results can only partially be compared with theirs, because, for instance, they also assessed callous-unemotional traits. In their study, only children who showed higher levels of conduct disorder but low levels of callous-unemotional traits had deficits in both affective and cognitive ToM. The children who showed higher levels of conduct disorder and callous-unemotional traits revealed relative competency in cognitive ToM, but deficits in affective ToM (Anastassiou-Hadjicharalambous & Warden, 2008). Another difference is that we integrated the cognitive and affective aspects of ToM into one latent factor, so we cannot differentiate children with high levels in only one of the two facets of ToM. Nevertheless, as our sample was a community sample, the probability that there was a substantial percentage of children with high levels of callous-unemotional traits is relatively low. For instance, Blair et al. (2014) report that callous-unemotional traits occur in fewer than half of those with conduct disorder, and the diagnosis of conduct disorder already requires serious clinical criteria. This is in contrast to the sample of Anastassiou-Hadjicharalambous and Warden (2008) that was preselected for children with conduct disorder identified in different programs for children with emotional and behavioral difficulties. This makes it relatively difficult to directly compare our results to theirs; however, probably the most similar group to our sample might be the one with conduct disorder, but low levels of callous-unemotional traits, who revealed deficits in both cognitive and affective ToM, which is in line with our result.

Furthermore, our result does not support the proposition by Happé and Frith (1996) about a ‘theory of nasty minds.’ We found no indication in our study that the children who display intact ToM might use their ability to infer others’ beliefs, intentions or feelings for antisocial behavior like, for instance, lying, cheating or teasing. However, we did not assess the application of ToM abilities in everyday life with items similar to the items of the Vineland Adaptive Behavior Scale, which Happé and Frith (1996) used in their study. This would be an interesting topic for future research. In addition, we found no indication for manipulating others' behavior, which Sutton et al. (1999a) proposed that children with higher ToM scores might show. However, this proposition cannot really be applied to our study, because we did not investigate bullying behavior.

Further, we also tested the relations in this study for moderation by gender and age and found no indication for it. This is an interesting result, because the boys in our sample had significantly higher conduct problem scores, and the girls showed significantly higher affective ToM scores. This is in line with previous research that has shown that boys usually display higher conduct problem scores than girls (e.g., Ravens-Sieberer et al., 2008), and that girls typically display superiority in ToM tasks compared to boys (e.g., Devine & Hughes, 2013). Different explanations can be found for the higher prevalence rate in boys for conduct problems. In the Dunedin study, most females who exhibited delinquent behavior only started after the age of 10, and thus would be in the adolescent-onset trajectory group (Moffitt & Caspi, 2001). Another longitudinal study found that most girls' conduct disorder started before the age of 10 (Keenan et al., 2010). However, the conduct problems in our study cannot be perfectly compared to the diagnosed conduct disorder and delinquent behaviors, which were noted in the two aforementioned studies. Altogether, many studies report higher prevalence rates for conduct problems in males at least before adolescence, and in adolescence prevalence rates are generally reported to increase. One reason that is frequently mentioned is the 'maturity gap,' which is felt by adolescents who feel that they are biologically mature, but not socially accepted as mature adults. This applies similarly to females and males and is one reason for the higher prevalence in conduct disorder symptoms in adolescence compared to early and middle childhood. However, in our study, age did not correlate with conduct problems, but with ToM. As explained before, an increase in ToM ability with age is expected and in line with previous research (e.g., Apperly et al., 2011). In sum, the results of Study 3 indicate that the link and underlying mechanisms between deficits in ToM and conduct problems seem to be similar for females and males, and among the three age groups of our sample (9.1 to 10.6 years, 10.6 to 11.6 years, 11.6 to 13.8 years).

To sum up the results considering the first research question, the reported findings of Studies 1 and 3 are in line with the SIP theory and the amygdala theory of violent behavior, respectively. Further, the first research question was answered as follows: Deficits in ToM are a risk factor for conduct problems, including different forms of aggressive behavior, physical and relational. This was true similarly for females and males, as well as three equally sized age-groups in middle childhood and early adolescence, which suggests that similar underlying mechanisms may be responsible for the observed relations. Naturally, fulfilling two of the basic requirements for establishing causality, temporal precedence and observed covariation, is not enough to determine a causal relationship (e.g., Valeri & VanderWeele, 2013). Among other factors, it is also necessary to eliminate plausible alternative explanations for an observed relationship (e.g., Preacher, 2015), which is very difficult to test, and cannot be tested sufficiently in this study.

The results of Studies 1 and 3 extend the previous evidence in different ways. First, the relations between ToM and different forms of aggressive behavior were examined longitudinally from middle childhood to early adolescence while controlling for stable between-person differences in aggression and for past levels of ToM, which has not been investigated in such a way before. This fosters a causal interpretation of the investigated paths (e.g., Hamaker et al., 2015). Second, reciprocal effects between deficits in ToM and aggression, including the reverse path from aggression to ToM, were investigated in this age range for the first time. Third, the link between ToM and conduct problems was examined during the transition from middle childhood to early adolescence, which extends the previously limited evidence in this age period. Fourth, we controlled ToM for verbal ability, which has not been done in all previous studies. Finally, gender and age differences were investigated in separate multi-group models in all analyses, which has also been neglected in prior research.

7.3. Deficits in EF as a risk factor for conduct problems

The second research question concerned deficits in EF as a risk factor for conduct problems in middle childhood and early adolescence. This was investigated in Studies 2 and 3, as well as in the Additional Analyses (see section 6.). In Studies 2 and 3, only cool EF processes were investigated, and thus when deficits in EF are discussed in this context, only cool EF processes are meant. In the Additional Analyses, also hot EF processes are discussed, and then the differentiation between cool and hot EF is mentioned.

In line with the SIP model, children with deficits in their social-cognitive processing are proposed to show more aggressive behavior than their peers (Crick & Dodge, 1994). Further, deficits in EF are related to impaired SIP in children who show higher levels of aggressive behavior (de Castro & van Dijk, 2018; Zadeh, Im-Bolter, & Cohen, 2007). This can also be explained in terms of impaired social problem-solving, which is defined as a function of EF, in children with higher levels of physical aggression (Séguin & Zelazo, 2005). Moffitt (1993) also proposed in her developmental taxonomy that children who show antisocial behavior from an early age display deficits in executive control functions, which are part of EF. This is in line with the frontal-lobe hypothesis of emotional and behavioral regulation that states that individuals who show higher levels of aggression, particularly physical aggression, display impairment in EF behavior (Séguin, 2009). Previous evidence on the association between deficits in EF and physical aggression and/or conduct problems comes from cross-sectional research (e.g., Schoemaker et al., 2013) and from research in early (e.g., Hughes & Ensor, 2008) and middle childhood (e.g., Eisenberg et al., 2004; Huang-Pollock et al., 2017), as well as from meta-analyses across the lifespan (Ogilvie et al., 2011;

Morgan & Lilienfeld, 2000). However, research on the relation between deficits in EF and relational aggression is more limited and mixed, and evidence on the relation to reactive and proactive aggression is also limited. Consequently, less is known on the relation between deficits in EF and specific forms and functions of aggression, and about this association in middle childhood and early adolescence. Further, the direction of effects is not clear, because most previous evidence was cross-sectional.

Consequently, in Study 2 deficits in EF were examined as a risk factor for physical, relational, reactive and proactive aggression longitudinally, which allowed us to control for earlier levels of aggressive behavior and to partial out between-person differences in aggression. Based on previous evidence, negative paths from EF at T1 to later physical (e.g., O'Toole, Monks, & Tsermentseli, 2017) and reactive aggression (Ellis, Weiss, & Lochman, 2009; Rathert, Fite, & Gaertner, 2011), and as a consequence of the mixed evidence a negative or positive path to relational aggression (e.g., McQuade, Murray-Close, Shoulberg, & Hoza, 2013), but no path to proactive aggression (Ellis et al., 2009) at T2 and T3 were proposed. We also investigated anger as a further underlying mechanism at T2. This is in line with the adapted SIP model by Lemerise & Arsenio (2000), which integrates emotions and emotional processes into the SIP model. Anger seems to play a particularly important role in the mediation between social-cognitive processes and aggressive behavior (de Castro et al., 2005), and deficits in EF are associated with anger (e.g., Bridgett et al., 2013). Consequently, anger was supposed to mediate the relation between deficits in EF and physical, relational and reactive aggression.

As a result in Study 2, the hypotheses concerning the negative paths from EF to later physical, and reactive aggression, as well as the proposed absence of a significant path to proactive aggression were confirmed. In addition, a negative path from earlier EF to later relational aggression was determined, which is also in line with our hypothesis. Consequently, children with deficits in EF at T1 at a mean age of 8 years showed higher levels of physical, relational, and reactive aggression at T2 at age 9 and at T3 at age 11, as rated by their teachers. Our results are controlled for age, gender, and information-processing capacity. We also included a random intercept for stable between-person differences in the forms and functions of aggression. According to Hamaker et al. (2015), this allows us to more accurately uncover causal relationships in within-person processes. Our results on the negative paths from EF to later physical and relational aggression extend cross-sectional and longitudinal findings from other age groups (e.g., Hughes & Ensor, 2008; Hughes et al., 1998; Ogilvie et al., 2011; Olson et al., 2017; Schoemaker et al., 2013). Our finding of the significant, negative paths from EF to later relational aggression confirms some of the cross-sectional evidence on middle childhood (McQuade et al., 2013), but it contradicts other research that found a positive link (Poland et al., 2016), and no link (Diamantopoulou, Rydell, Thorell, & Bohlin, 2007b). The reason for these contrasting previous results could be that relational aggression can take different forms, ranging from relatively simple, direct forms (e.g., threatening a friend with ending the friendship) to more complex, indirect forms (e.g., mobilizing peer group members against a child that should feel excluded) (Crick et al., 1999). Naturally, the latter forms require higher socio-cognitive skills, and thus could be unrelated or even positively related to relationally aggressive behavior in contrast to the former, which might be related to deficits in EF and thus show a negative relation to EF. Our items mainly covered more direct forms of relationally aggressive behaviors, which explains the negative relation between EF and relational aggression. Consequently, it could be necessary to use a more fine-grained assessment of relational aggression that considers the complexity of relationally aggressive behavior and includes more than three items for each form of aggression, as in our study.

Discussion

Our results can be interpreted in line with the SIP model, as children who had deficits in their cognitive processing in social situations, such as in impaired inhibition of prepotent or dominant responses that were inadequate social responses, showed more physically, relationally, and reactively aggressive behavior about one and two years later. Further, our results can be interpreted in terms of impaired social problem-solving, such as failure to plan a solution (Séguin & Zelazo, 2005), in the children that showed later higher levels of physical, relational, and reactive aggression. This extends the concept of social problem-solving, which has until now mainly been considered for physical aggression. Our results are also in line with the frontal-lobe hypothesis of behavioral and emotional regulation (Séguin, 2009), which states that children who behave aggressively have deficits in their regulation of emotions and behavior, but also extends the empirical results to relational and reactive aggression. The nonsignificant path between EF and proactive aggression is also consistent with the view previous cross-sectional studies (Ellis et al., 2009; White, Jarrett, & Ollendick, 2013) regarding this association that did not find a significant path between EF and proactive aggression. It is also in line with research on deficits in EF in the association with callous-unemotional traits, which are conceptually linked to proactive aggression (e.g., Pardini et al., 2003). These traits have been found to be unrelated to deficits in EF (Mitchell, Colledge, Leonard, & Blair, 2002), or to rather superior levels in processes related to EF (Tye et al., 2017). Further, our findings support the theoretical and empirical differentiation between the two functions of aggression, because higher levels of EF enable children to plan ahead and behave in a more sophisticated way, which is characteristic of proactive aggression. In contrast to reactive aggression that refers to more impulsive acts that do not require planning, and thus the inability to plan and inhibit behavioral impulses, explains the negative path between EF and reactive aggression one or two years later.

Discussion

Considering the mediational effect of anger, this was only confirmed for physical aggression. Our results indicate that children with deficits in EF at T1, such as in their shifting between different response sets according to the situation at hand, displayed more parent-rated anger about one year later at T2, which predicted more teacher-rated physical aggression at T3. The size of the mediational effect was small; however, our study was the first – as far as we know – to uncover this effect over a time span of three years in middle childhood. By contrast, the mediation between deficits in EF and relational aggression by anger was not confirmed, although the path from anger to later relational aggression only narrowly missed the level of significance in our study. Therefore, this tentatively suggests that anger could also be involved in the mediation of the association between deficits in EF and relational aggression. For reactive aggression, there was no mediational effect of anger in the path from deficits in EF to reactive aggression. Deficits in EF were a predictor of higher levels of anger, which is in line with previous evidence (e.g., Carlson & Wang, 2007), such as from research on children with irritability – a chronic tendency to feel anger (e.g., Perlman et al., 2015) – although anger was not related to later reactive aggression. This was surprising, because anger and irritability are both proposed to be an important impelling factor in reactive aggression (e.g., Arsenio et al., 2000; Eisenberg et al., 1999; Leibenluft & Stoddard, 2013; Olson et al., 2005). One explanation could be that other moderating variables that were not investigated are important in this relation, for instance, less voluntary and more reactive aspects of control or anger regulation (Rothbart et al., 2001), or other cognitive information-processing variables like hostile attribution bias (de Castro, Veerman, Koops, Bosch, & Monshouwer, 2002) or ToM (Renouf, Brendgen, Séguin, et al., 2010). Another reason could be that previous evidence mainly comes from studies on the relations between chronic anger, irritability, and disruptive symptoms in clinical samples, with no differentiation between reactive and proactive aggression (e.g., Wakschlag et al., 2015).

Discussion

Considering the direction of effects, we found evidence for earlier deficits in EF predicting later, higher levels of aggressive behavior. However, we could not examine the reverse direction in our study, because the stability of EF was so high from T1 to T2 that we could not include EF at T2 in our model.

In further analyses, moderation of all paths by gender and age was examined and not found. However, we found gender differences in some study variables, particularly in physical aggression, in which boys scored higher than girls at all three time points. This is in line with previous research, including a meta-analysis (Card et al., 2008). Further, girls showed higher cognitive flexibility, inhibition, and planning ability. The evidence regarding gender differences in EF is mixed and limited. Some studies found no gender difference (Huang-Pollock et al., 2017) and others found males to perform better on tasks requiring speed (Brocki & Bohlin, 2004), whereas some found evidence for girls performing better on tasks involving control, such as effortful control (e.g., Olson et al., 2005). However, in the last study they also did not find a moderation in the relation between effortful control and aggressive behavior in children (Olson et al., 2011), although this was only rarely investigated. Altogether, in our study we found no indication for gender differences in the paths between EF, anger and the aggressive behaviors. Regarding the influence of age in our study, we found a significant positive correlation between age and EF, indicating that older children showed higher EF abilities. This finding is in line with previous research that found ongoing development of EF from childhood through adolescence (e.g., Blakemore & Choudhury, 2006; Zelazo & Carlson, 2012). Thereby, each facet of EF follows a different developmental course, including regressive and progressive phases, some showing a more protracted developmental trajectory, for instance set-shifting or cognitive flexibility, as it is also called (e.g., Best & Miller, 2010). However, in this study the single facets of EF were included in one latent factor; therefore, conclusions regarding their specific development cannot be drawn. In sum, our results on gender and age differences suggest that the processes and underlying mechanisms in the paths from deficits in EF to aggressive behavior might be similar for girls and boys and for the three different age groups in our sample from age 8 to 11 years.

As a conclusion, the findings of Study 2 can be interpreted as an answer to the second research question, partially confirming that deficits in EF are a risk factor for conduct problems that include physical, relational, and reactive aggression on an intrapersonal level from middle childhood through early adolescence. Anger seems to play a mediating effect in the relation between deficits in EF and physical aggression, as well as possibly relational aggression.

In Study 3, the relation between deficits in EF and conduct problems was investigated cross-sectionally only with data from T3. However, because of the large age range in our sample (9 to 13 years), the study connected the developmental periods of middle childhood and early adolescence. Previous evidence on the association between deficits in EF and conduct problems – until now – has focused mainly on early (e.g., Schoemaker et al., 2013) and middle childhood (e.g., Eisenberg et al., 2004). Research on adolescence is limited to older studies on incarcerated and convicted adolescents (see Morgan & Lilienfeld, 2000), with the exception of one study (Dolan & Lennox, 2013). Based on previous evidence and theoretical considerations, such as the problem-solving account (e.g., Séguin & Zelazo, 2005) and the frontal lobe hypothesis (e.g., Séguin, 2009), we hypothesized a direct negative path from EF to conduct problems. In addition, we examined a possible negative indirect path from EF to conduct problems through ToM. Such an indirect link was suggested by Hughes et al. (1998) and is discussed in the next paragraph (7.4). According to our results, in Study 3 we found no direct link between EF and conduct problems, only an indirect link through ToM. This means that only children who had deficits in both EF and ToM showed more conduct problems. Therefore, the hypothesis of a negative path from EF to conduct problems was not confirmed considering the direct path. Therefore, we could not confirm the social problem-solving account (e.g., Séguin & Zelazo, 2005) or the frontal lobe hypothesis (e.g., Séguin, 2009). Different reasons could be responsible for the missing link, including differences in participants' age or clinical status, as we investigated a non-clinical community sample because Schoemaker et al. (2013) found in their meta-analysis a stronger relation between deficits in EF and conduct problems for clinically referred samples. It could also be a consequence of different measurement methods for conduct problems in contrast to the diagnosis of conduct disorder, or for measuring the facets of EF. The topic of measuring EF is very complicated as a consequence of the ongoing development of EF throughout childhood and adolescence and the consequent necessity to avoid ceiling or floor effects in the measurement. Further, a differentiation is made between simple and complex tasks, which makes it even more difficult to compare different tasks to measure the same facet of EF (e.g., Best & Miller, 2010). Another reason could be that we investigated the impact of EF and ToM in tandem in our analysis, which will be discussed in the next section (7.4), although we found no significant path from EF to conduct problems when examining EF as the sole predictor ($\beta = -.09$, $p = .11$). However, other research also failed to find a significant effect of EF on conduct problems (Hughes & Ensor, 2006). Furthermore, an earlier quantitative review also found no relation between EF and conduct disorder (Pennington & Ozonoff, 1996). And, Morgan and Lilienfeld (2000) argued that, because children with conduct problems display deficits in neuropsychological tasks in general, EF deficits could be a possible consequence, but not a sufficient proposition for conduct problems. However, this has to be investigated in future studies, including different neuropsychological assessments.

In Study 3, moderation by gender and age was also examined, but not found. This is noteworthy, because the boys in our sample had significantly higher conduct problem scores and the girls showed significantly higher set-shifting and inhibition ability. As mentioned before, evidence on gender differences in EF is limited and mixed, although some studies found girls' superiority on tasks involving inhibitory control (e.g., Berlin & Bohlin, 2002), which might have played a role in our set-shifting, as well as in the inhibition task. By contrast, other research found boys performing better on tasks requiring speed (e.g., Brocki & Bohlin, 2004). These contrasting results emphasize the need to further investigate gender differences in EF, particularly the relation to aggressive behavior. The few studies that investigated gender differences in the association with conduct problems also found contrasting results. Two studies found that boys with lower levels of EF showed more conduct problems (Raaijmakers et al., 2008; Lonigan et al., 2017), whereas the study by Huang-Pollock et al. (2017) found no differences in the relation between working memory and externalizing behavior disorders. Similarly, we found no indication for moderation by gender in our paths. This suggests that similar processes may underlie the effects in our sample across females and males. Considering age differences, these were also very rarely investigated in previous research on EF and conduct problems. We found a significant, positive correlation between age and EF, but no indication for moderation by the three age groups of our sample (9.1 to 10.6 years, 10.6 to 11.6 years, 11.6 to 13.8 years) in our paths. As explained before, the correlation by age and EF is in line with research indicating ongoing development of EF throughout middle childhood and adolescence (e.g., Best & Miller, 2010). However, the processes underlying the (non-significant) path from EF to conduct problems seem to be similar across the age range from 9 to 13 years during the transition from middle childhood to early adolescence, although this has to be investigated further in future studies.

In the Additional Analyses (section 6.), the growth and prediction of conduct problems over the course of our study were examined. Considering the growth of conduct problems, we found a very low baseline level at the beginning of our study with small interindividual variability, and a small, but significant and linear decrease over the course of our study with no interindividual differences. We expected an increase in conduct problems over the course of our study based on the propositions of the developmental taxonomy by Moffitt (1993), for instance, that in adolescence the maturity gap is a reason why more children and adolescents show conduct problems. Different reasons could be responsible for why we did not encounter an increase in conduct problems over the course of our study. One reason is that the children in our study only approached the transition to early adolescence at the third and last measurement point (mean age 11 years). So, it could be that the majority of them did not yet feel the discomfort of the maturity gap, which is assumed to be one reason for the increase in conduct problems in adolescence. A second reason is that the conduct problem scores were at a very low level in our sample, and so floor effects make it difficult to detect any change within them.

Regarding the prediction of conduct problems, only deficits in emotional and inhibitory control and planning were related to baseline levels of conduct problems, but not behaviorally measured cool EF with the three facets of working memory, inhibition, and set-shifting. The result that deficits in emotional and inhibitory control were related to higher levels of conduct problems is in line with propositions by Frick and Morris (2004). They argue that children who show deficits in regulating their emotions and behavior tend to experience more negative emotions, which – in combination with their difficulties controlling themselves in emotionally arousing situations – hinders the application of socio-cognitive skills. This is also in accordance with the adapted SIP model by Lemerise and Arsenio (2000) that states that emotional processes, including regulating processes, mediate between socio-cognitive processes, the behavior in the situation, and the social knowledge. Inhibitory control is also a subdomain of effortful control, which has been negatively related to conduct problems in children (see Eisenberg et al., 2010). Similarly to emotional control processes, inhibitory control processes can also be conceptualized as being part of hot EF (e.g., Zelazo & Carlson, 2012); however, the questionnaires that we used in our study were not typical hot EF procedures, so they cannot be easily compared with research on hot EF. Therefore, we compared our results with research on deficits in emotional, inhibitory and effortful control that also found a link to higher levels of conduct problems (for a review, see Eisenberg, Spinrad, & Eggum, 2010). Further, planning that is also conceptualized as a facet of cool EF showed a negative path to the baseline level of conduct problems, which confirmed our assumption. This result is in line with the problem-solving account (Sèguin & Zelazo, 2005), in which deficits in planning hinder children's abilities to solve a social problem. Moreover, we did not find a significant path from cool EF to the baseline level of conduct problems, so our hypothesis on a negative path was not confirmed. Different reasons could be responsible for the missing link, for instance, all other variables were measured with questionnaires that were rated by parents or, considering planning, by teachers, so the overlap between them might have masked the association between the behavioral measures of cool EF and conduct problems. However, not all previous research has distinguished between cool and hot EF processes, such as a meta-analysis that found a strong relation between conduct disorder and problems and EF (Ogilvie, Stewart, Chan, & Shum, 2011). So, the possibly stronger relation between hot EF processes and conduct problems could have masked a weaker or nonsignificant link between cool EF and conduct problems. Similarly, Séguin (2009) proposes that the relation between physical aggression, as one part of conduct problems, and hot EF might be stronger compared to cool EF. Nevertheless, Zelazo et al. (2010) argue that cool and hot EF always work together, so the missing link could as well be the consequence of the different sources of our data and the different measurement methods in our study. Another issue is that, particularly in children, it is often difficult to clearly distinguish between cool and hot EF processes, because every test situation might involve some motivational aspects for them (e.g., Welsh & Peterson, 2014). Altogether, the reason why we did not find a link from deficits in cool EF to conduct problems has to be solved in future studies that include multi-informant and multi-method measurements. A last question in the Additional Analyses was whether the developmental course of conduct problems can be predicted by emotional and inhibitory control, planning, and cool EF. The result was that the developmental trajectory, that is, the growth, of conduct problems could not be predicted by the variables of our study. Possible reasons could be the restricted variance in conduct problems as a consequence of the floor effects, or other moderating variables like parenting practices that were not investigated in our study. Altogether, deficits in emotional and inhibitory control processes, as well as deficits in planning, seem to play an important role in the prediction of conduct problems at a mean age of 8 years.

To sum up the results in light of the second research question, it could be possible that deficits in cool EF are only a risk factor for different kinds of aggressive behavior, physically, relationally, and reactively, as found in Study 2, but not for conduct problems in general. At least, this could not be confirmed in the age range from 6 to 13 years in the Additional Analyses. However, it has to be noted that in Study 3 the latent factor for EF had three indicators, working memory, inhibition, and set-shifting, whereas in Study 2, planning was a fourth indicator for EF. This could also be a reason for the differing results. Deficits in planning are seen as a particularly important impelling factor in solving social problems (e.g., Séguin & Zelazo, 2005), and were also directly related to conduct problems in the Additional Analyses. The reasons for excluding planning from Study 3 were primarily methodological, because the model with planning as a fourth indicator of EF or as an additional factor did not result in a solution, and thus these models seemed not to appropriately fit the data.

Consequently, the second research question was answered as follows: Deficits in cool EF serve as a risk factor for later physical, relational, and reactive aggression over the course of middle childhood throughout early adolescence, and deficits in emotional and inhibitory control, as well as in planning, serve as a risk factor for conduct problems in middle childhood. We could not confirm that deficits in cool EF, apart from planning, are a direct risk factor for conduct problems, at least in our community sample with a mean age of 8 years. The results extend previous evidence in two ways. First, we determined that relational and reactive aggression – apart from physical aggression – should also be considered as possible outcomes of deficits in cool EF, which has been neglected in previous research. Second, we found evidence that deficits in cool EF are a risk factor for physical, relational, and reactive aggression over a period of three years on an intrapersonal level, which goes beyond cross-sectional and correlational results. By including a random intercept, we controlled for trait stability of aggression, which allows the disentangling of within-person processes from stable between-person differences, according to Hamaker et al. (2015). This is noted as a better way to reveal causal relations between different variables. In Study 2, two requirements for causality were fulfilled, namely observed covariation and temporal precedence (e.g., Valeri & VanderWeele, 2013). Third, we explored whether the longitudinal relations between deficits in cool EF and forms and functions of aggressive behavior, as well as the cross-sectional relations between cool EF and conduct problems, were moderated by gender. As a result, we found no indication for moderation, and consequently, similar effects seem to underlie these paths in females and males. Fourth, we also examined whether the cross-sectional relations between cool EF and conduct problems were moderated by age. We found no indication for differences in the investigated paths between three different age groups within a range from 9 to 13 years.

7.4. The study of deficits in EF and ToM in tandem as risk factors for conduct problems

Most research has until now considered deficits in EF and ToM as risk factors for conduct problems in separate analyses or studies, despite the evidence for their interrelatedness throughout development (e.g., Devine & Hughes, 2014) and evidence for a dependence of ToM performance and competence on EF (e.g., Apperly et al., 2011). Research on younger children found mixed evidence for cross-sectional relations, when including EF and ToM in the same analysis, with some evidence for a direct effect of deficits in EF on conduct problems (e.g., O'Toole et al., 2017) and some against it (e.g., Hughes & Ensor, 2006). Moreover, one study suggested an indirect link between deficits in EF and conduct problems through impaired ToM in addition to a direct link (Hughes et al., 1998). As a consequence, the third research aim was to determine whether deficits in EF and ToM are risk factors for conduct problems when they are studied in tandem during the transition from middle childhood to early adolescence. This was investigated in Study 3.

In light of previous evidence and theoretical considerations, as explained before, we proposed that deficits in both EF and ToM are risk factors for conduct problems when they are studied in tandem. Further, we examined whether an indirect link from deficits in EF to conduct problems through impaired ToM would be present. Our hypotheses were partially confirmed by a negative direct path from ToM to conduct problems (as discussed in 7.2), a positive direct path from EF to ToM, and a negative indirect path from EF to conduct problems through impaired ToM. However, we found no direct path from deficits in EF to conduct problems (as discussed in 7.3).

The direct positive path from EF to ToM supports the idea of a dependence of ToM deployment on EF (e.g., Apperly et al., 2011), because children in our study who showed lower levels of EF also showed lower levels of ToM. However, as our study is only cross-sectional, this cannot be interpreted as a causal effect. Nevertheless, our finding is in line with previous evidence on this link in middle childhood (e.g., Austin et al., 2014; Bock et al., 2014; Wang et al., 2016) and initial evidence from adolescence (Vetter et al., 2013).

Discussion

The absence of a negative direct path from EF to conduct problems could theoretically be a consequence of investigating deficits in EF and ToM as risk factors for conduct problems in tandem. However, we also found no direct path from EF to conduct problems when examined alone in this study. So this is not an explanation for the missing link, and other reasons must be responsible (as discussed in 7.3). However, we did find a negative indirect path from deficits in EF to conduct problems through impaired ToM. As a consequence, children who displayed deficits in their EF, for instance, in inhibiting prepotent behavioral impulses and inferring the affective and cognitive mental states of others, exhibited higher levels of conduct problems compared to other children in our study. So, this suggests that deficits in EF alone are not sufficient to explain the conduct problems of some children, at least in our community samples with an age range from 9 to 13 years. Our finding is in line with another study on younger children that found that deficits in EF and deficits in ToM were associated with conduct problems in a group of children only compared to a control group of typically developing children (Hughes et al., 1998), which also suggests an indirect link between deficits in EF and conduct problems through impaired ToM.

In addition, we also investigated gender and age differences in the indirect path from EF through ToM to conduct problems, and in the path from EF to ToM. We found no indication for moderation by gender or age group in these paths. However, we found differences between females and males in some of the study variables, in particular a higher frequency of conduct problems in boys and female superiority in affective ToM, set-shifting, and inhibition (as discussed in 7.2., 7.3). Further, age correlated positively with both ToM and EF, which indicates that older children displayed higher ToM and EF ability. Nevertheless, in the investigated paths of our model, we found no indication for moderation by either gender or age group. Consequently, this suggests that similar processes underlie these paths for females and males and the three age groups (9.1 to 10.6 years, 10.6 to 11.6 years, 11.6 to 13.8 years).

Taken together, the answer to the third research question is as follows: When studying deficits in EF and ToM in tandem, cool EF seems to be a risk factor only through deficits in ToM for conduct problems, but not the sole risk factor, at least in our non-clinical community sample in the transition from middle childhood to early adolescence. However, we did not investigate hot EF in this context. The result extends previous research in three ways. First, we investigated deficits in EF and ToM as risk factors in tandem and in one analysis, which has not been done often in previous research in this age range. Second, we connected research on middle childhood and early adolescence in our study. Finally, we explored moderation by gender and age in our results.

7.5. Strengths and limitations of the presented studies

This thesis has several strengths. First, in two longitudinal studies, Studies 1 and 2, aggressive behavior, ToM, and EF have been controlled for earlier levels of the same construct (construct stability), which fosters a causal interpretation of the investigated paths (e.g., Marmor & Montemayor, 1977). Second, the forms and functions of aggressive behavior have been controlled for stable between-person differences, which helps to separate the within-person from the between-person processes (Hamaker et al., 2015), and consequently allows for a more accurate investigation of causal relations. Third, the reciprocal relations between two forms of aggressive behavior and ToM were examined, which has not been done in previous research apart from one exception (Song et al., 2016), but not in this age period. Fourth, we controlled ToM for verbal ability and EF for information-processing capacity, because both are closely related and need to be controlled for the respective influence of the other. Fifth, we investigated the cognitive and affective facets of ToM. Sixth, moderation by gender and age group was explored in separate multi-group models. Seventh, relational and reactive aggression – apart from physical aggression – were also considered as possible outcomes of deficits in cool EF, which has been neglected in previous research. Eighth, we investigated a mediation by anger in the relation between EF and physically, relationally, and reactively aggressive behavior over the course of three years, which has not been done before. Eighth, the impact of deficits in EF and ToM were examined in tandem in one analysis, which has not been done very often in this age range during the transition from middle childhood to early adolescence. And, finally, the large community sample, which can also be seen as a limitation, has to be mentioned as an advantage, because some of the research questions and perspectives have not been investigated in such a large community sample before.

At the same time, some limitations of the presented studies have to be noted. First, the participants of the studies came from selected schools in Brandenburg and a variety of rural and urban areas. However, as the mean levels of conduct problems and aggressive behaviors display, the participating children presented very few conduct problem behaviors or aggressive behaviors. This floor effect might have restricted the variance in our analysis, and could be another reason why, for instance, we did not find an effect of deficits in EF on conduct problems. Further, as Hughes and Ensor (2008) have shown, social disadvantage is positively correlated with conduct problems, and negatively correlated with EF and ToM. As the proxy for SES in our studies shows, we had very few socially disadvantaged families, as the occupational level of only 11% of mothers and 15% of fathers was classified as unemployed or as holding a job that could not be classified. This also displays that our proxy for SES was not very exact in measuring socio-economic level or status, and we did not note any other markers of social disadvantage, such as whether the family lived in council accommodation or in a very poor neighborhood, like Hughes and Ensor (2008) did in their study. Altogether, it seems that the selected schools probably did not represent enough variety in terms of socially disadvantaged families, which might have limited the results, particularly in Study 3 and in the additional results.

Second, the ToM measures are another reason that might have limited the results, but in the other direction. The cartoons for cognitive and affective ToM that were used to measure ToM at T1 and T2 were a selection of cartoons from the study by Sebastian et al., (2012). They seemed to be too easy for our children, at least at T2, where we encountered ceiling effects, particularly in the cartoons for cognitive ToM. Sebastian et al. (2012) used these cartoons in a neuroimaging study, and they may be better suited for activating parts of the brain than detecting individual differences in inferring mental states in children, who by their age should have already acquired second-order false belief. The cartoons were used in Study 1, and the ceiling effects might have limited our results, because they reduced the variance in ToM and made it harder to detect very skilled children and interindividual differences between them, as well as to identify potential relations to other variables. That is the reason why we changed the measurement method for ToM at T3. Instead of the cartoons, we used one task for cognitive ToM, an adapted version of the Belief-Desire-Reasoning Task (Apperly et al., 2011), and an adapted version of the Faces Task (Vetter et al., 2013). The Faces Task showed good variability apart from two items, for which more than 90% of the children chose the correct response, and which displayed ceiling effects and were excluded from the study. However, the Belief-Desire-Reasoning Task also seemed a bit too easy for our children, because across all items with a categorical coding of 0 or 1, the mean level for cognitive ToM was $M = 0.80$. These ceiling effects might also have limited our results in Study 3. Perhaps another measurement method with a higher difficulty and also higher environmental validity would have been more appropriate, such as the Silent Films Task by Devine and Hughes (Devine & Hughes, 2013), which is also used to measure cognitive ToM in middle childhood.

In sum, two of the ToM measures might have been too easy, and as a consequence ceiling effects might have limited the results of Studies 1 and 3.

A third limitation is the lack of investigation of attention deficit and hyperactivity symptoms (ADHD). These symptoms show high comorbidity with conduct problems and it has been shown that children, mostly boys, follow a trajectory from ADHD in early school years to oppositional defiant disorder and subsequent conduct disorder (e.g., Burke et al., 2005). Further, as Pennington and Ozonoff (1996) argue, EF deficits could be specific to children who show conduct problems and comorbid ADHD symptoms. And, other studies have shown differential effects in children with comorbid ADHD symptoms and conduct problems (e.g., Antonini, Becker, Tamm, & Epstein, 2015). As a consequence, including ADHD symptoms in Study 3 might have led to different results. The primary reason why they were not included in Study 3 was an overlap with another PhD project that was also part of the PIER study. So, in sum, the inclusion of ADHD symptoms would probably have been a beneficial addition to Study 3.

Fourth, another limitation lies in the measurement of hot EF, because according to Séguin (2009), children who show higher levels of aggressive behavior also might not only show deficits in cool EF, but also in hot EF, particularly in measures of decision making. Also, Blair et al. (2014) note that deficits in decision making are a common neurocognitive impairment in children with conduct problems. Further, hot and cool measures of EF are supposed to operate together as part of a more general adaptive function (Zelazo & Carlson, 2012). Although in the additional results emotional control and inhibitory control were examined, which could be related to hot EF measures, as predictors of conduct problems, however, these are not measures of decision making. There was a shortened child version of the Iowa Gambling Task in our study, but models that included it did not show an adequate fit, and this measure was also part of another PhD project of the PIER Study, so it had to be excluded from the present thesis. However, generally the inclusion of hot EF measures, particularly of decision making, might have been a valuable addition to the presented studies, for instance to Studies 2 and 3.

A fifth limitation is the source for measuring conduct problems and aggressive behaviors in our studies. Study 1 used teacher reports for physical and relational aggression, as did Study 2 with the addition of teacher-reported reactive and proactive aggression and parent reports of habitual anger. Study 3 analyzed conduct problems rated by children's parents on a selection of CBCL items. In the Additional Analyses, parent-rated conduct problem symptoms were also used, but from the SDQ. Moffitt et al. (2001) and Hughes and Ensor (2008) suggest that multi-informant ratings are probably the most accurate way of measuring conduct problems for different reasons. First, with increasing age, a greater amount of time is spent with peers. So, parents and teachers are not able to accurately report all conduct behavioral symptoms. Second, both have different interests in either reporting or not reporting conduct problem symptoms. As a consequence, Moffitt et al. (2001) suggest using multi-informant ratings including parents, teachers, and also neutral external ratings of, for instance, video-taped sequences. They also propose the inclusion of self-reports and criminal conviction data starting at the beginning of middle childhood and early adolescence. Altogether, the inclusion of a self-report, as well as measures by parents and teachers of aggressive behaviors and other conduct problem symptoms, which could be aggregated into a measure for conduct problems, might have been a feasible alternative for the longitudinal study as a whole.

A final limitation is that we did not assess SIP specifically despite using the SIP model as a theoretical background for the thesis. Other studies used vignettes or semi-structured interviews to measure interindividual differences in specific social-information processes (e.g., de Castro et al., 2005). Previous research confirmed relations between deficits in ToM and deficits in SIP (e.g., Choe et al., 2013), so we assumed that the deficits in ToM in our children are related to deficits in SIP; albeit, we did not test this assumption empirically.

In sum, different limitations are noted regarding the selection of schools and the proportion of socially disadvantaged families, the ToM measures, the inclusion of ADHD symptoms and hot EF measures, particularly decision making, the use of multi-informant measures for conduct problems, and measures for specific SIP. Despite these limitations, this thesis extends the previous research and highlights the need to consider reciprocal relations between ToM and different forms of aggressive behavior. Future research should extend this transactional focus to include research on ToM, EF, and conduct problems, including different forms and functions of aggression. The results from such longitudinal studies are important for tailoring programs to prevent conduct problems in children.

7.6. Implications for future research and practice

The SIP model served as a theoretical background for this thesis (Crick & Dodge, 1994) to explain aggressive behavior, and the amygdala theory of violent behavior (e.g., Blair et al., 2014) was used to explain violent and antisocial behavior in children with conduct problems. The first research question was whether deficits in ToM are a risk factor for conduct problems, including different forms of aggressive behavior, from middle childhood through early adolescence. This was confirmed for physical and relational aggression and other conduct behavioral symptoms, such as violent and antisocial behavior. This suggests that deficits in ToM play a decisive role in the development and maintenance of conduct problems. Further, our results also suggest that deficits in ToM are involved in transactional processes with aggressive behavior, because we also found the reverse path from relationally aggressive behavior to deficits in ToM. This is in line with the propositions of the SIP model and the ontogenic processes model of externalizing behavior (Beauchaine et al., 2015). These findings argue for the proposition that deficits in ToM are related to deficits in SIP (e.g. Choe et al., 2013), and that deficits in ToM might be a valuable addition to the SIP model.

Empirically, deficits in ToM have already been linked to deficits in SIP, such as in attributing hostile intent (Step 2; Choe et al., 2013), forming a mental representation of a social situation (Steps 1 and 2; Teufel et al., 2010), and adjusting a response decision (Step 5; Grueneisen et al., 2015). Deficits in ToM could be a part of the specific individual socio-cognitive capacities that a child brings into a social situation, for instance in the form of a 'database' or 'latent mental structures' as they were called by Crick and Dodge (1994), or "offline factors" as they are called by de Castro and van Dijk (2018). Crick and Dodge (1994) also differentiated between "on-line processing" (sic, p. 78) and latent mental structures. They proposed that the latter influence a child's on-line processing, and that this directly influences a child's social behavior in a social situation (Crick & Dodge, 1994). In the present thesis, no measures for specific SIP processes were included; therefore, this is only a theoretical assumption. However, deficits in ToM were attributed to higher levels of conduct problems, as well as to higher levels of future physically and relationally aggressive behavior, and in turn the latter impacted future deficits in ToM in our children. Consequently, it could be assumed that deficits in ToM play an important role in the development and maintenance of aggressive behavior. It should be the aim of future studies, to test whether deficits in ToM are a valuable addition to SIP theory in explaining aggressive behavior in children, and how deficits in ToM and social-information processes interact throughout development. This could be done by measuring specific social-information processes, for instance with vignettes, along with deficits in ToM and aggressive behavior in the same sample. It might then be determined whether or not deficits in ToM explain significant further variance in aggressive behavior.

To explain the relation between deficits in ToM and violent and antisocial behavior in the children, we used the amygdala theory of violent behavior (e.g., Blair et al., 2014). It proposes that ToM and empathy are closely related, and that deficits in both are also related (e.g., Shamay-Tsoory et al., 2003). Further, ToM is assumed to be a prerequisite to empathic responding, which is proposed to be involved in an important mechanism inhibiting violence and antisocial behavior. These propositions were originally formulated to explain the antisocial and violent behavior of children who display conduct disorder with callous-unemotional traits (e.g., Blair et al., 2014; Marsh et al., 2008). However, Sharp (2008) also explained the link between deficits in ToM and violent behavior in a community sample of children with conduct problems in reference to deficits in the violence-inhibiting mechanism. Another line of theory argues that the ability to take another's perspective, an important aspect of ToM, is essential for intentional moral conduct (e.g., Hoffman, 1987). Therefore, it is assumed that deficits in ToM are implicated in conduct problems (e.g., Anastassiou-Hadjicharalambous & Warden, 2008; Hare, 1970). These two theoretical lines of argumentation have the common feature of involving deviant functioning of the amygdala. As a consequence, it could be that deviant functioning of the amygdala is not only present in children who show conduct problems combined with callous-unemotional traits, as has been shown by Sebastian, McCrory, Cecil, et al. (2012), but also in non-clinical samples with conduct problems and no indication of callous-unemotional traits. However, this is also a topic for future research that includes neuroimaging techniques to examine whether children from community samples who show symptoms of conduct problems also display deviant functioning of the amygdala, and also whether or not this is correlated with callous-unemotional traits. Similarly, research should continue to distinguish deficits in empathy and ToM from one another, as Shamay-Tsoory et al. (e.g., 2003) have already shown. It would be important to know whether the differentiation between deficits in empathy and ToM is essential for studying conduct problems in particular, or whether the overlap between the two is the decisive element for etiological models of conduct problems.

Another point for future research is the question to what extent the identified cognitive deficits in ToM and EF of children who show higher levels of conduct problems interact with environmental factors, such as coercive parenting behavior, which has also been linked to externalizing behavior (Snyder, 2015). Such complex longitudinal interactions between vulnerabilities on the individual level and interdependent contextual risk factors have been proposed as decisive for the development of externalizing psychopathology (Beauchaine et al., 2015). Similarly, Fairchild et al. (2013) determined that the quality of children's early environment is supposed to moderate between individual vulnerabilities and the age-of-onset of antisocial behavior. Thereby, Moffitt's (1993) developmental taxonomy is integrated into a more recent perspective, in which it is proposed that all children with conduct problems display certain individual neurocognitive vulnerabilities, such as deficits in EF, independent of the age when they first displayed symptoms of conduct disorder (i.e., early onset vs. adolescence-onset). It has to be tested in future research whether this assumption can be confirmed for deficits in EF, because in the present thesis deficits in EF were only predictive of different kinds of aggressive behaviors and not of conduct problems in general. Only deficits in planning and emotional and inhibitory control predicted conduct problems in our study. It is an interesting question whether different facets of EF predict early vs. late onset of conduct problems; however, we did not have information on the age of onset in our study, so this is also a topic for future research.

Furthermore, in terms of equifinality and multifinality (e.g., Gatzke-Kopp et al., 2012), it would be necessary to examine whether children with deficits in ToM and/or in EF display a high risk of developing only conduct problems or also comorbid ADHD symptoms, as well as other psychopathologies. According to the transactional perspective of the ontogenic processes model of externalizing behavior, externalizing disorders including conduct problems are more accurately viewed as a spectrum of related behavioral syndromes than as discrete diagnostic categories (Beauchaine & Hinshaw, 2015). Therefore, more longitudinal research would be needed to determine which combinations of environmental factors together with deficits in ToM and/or in EF are most likely to predict conduct problems and subsequent conduct disorder, and whether they also predict comorbid ADHD symptoms, or other externalizing disorders, for example. As a consequence, research on different externalizing disorders should be combined to identify developmental risk factors that lead to any of the externalizing behavior disorders.

An important implication of the presented results for practice is their application to prevent conduct problems in children. One problem in preventing conduct problems is that many interventions that target specific diagnoses of externalizing disorders are less effective than they should be (e.g., Beauchaine & Hinshaw, 2015). Furthermore, a comprehensive understanding of the etiology of conduct problems is a prerequisite of maximally effective interventions (Preskorn & Baker, 2002). Therefore, new research results should be integrated into treatment strategies for conduct problems. Different programs for the prevention of conduct problems that are evidence-based and informed by recent research results exist already (e.g., The SMART-program; August, Piehler, & Bloomquist, 2016). They include a developmental life course orientation that identifies personal and contextual characteristics associated with intraindividual vulnerabilities. Similarly, also deficits in ToM and/or in EF as intraindividual vulnerabilities should be integrated as target aims in interventions preventing conduct problems. Other programs that focus on cognitive variables, such as the 'Promoting Alternative Thinking Strategies' program (PATHS; Crean & Johnson, 2013), have already noted success in reducing the risk of aggressive problem solving in elementary school students. Further, different interventions to promote the development of EF have been proven to be successful in improving EF in children between the ages of 4 and 12 years (Diamond & Lee, 2011). Thereby, it has been shown that children with initially worse EF benefit the most, and it is recommended not to focus narrowly on improving EF, but to also address emotional, social, and physical development. Furthermore, programs to foster ToM development are also beginning to be developed, like promoting conversations about mental states and others' emotions in classrooms. Lecce and colleagues have shown that with a simple to administer training program focused on fostering ToM abilities through conversations about mental states, children exhibited greater gains in ToM development than a control group in middle childhood (Lecce, Bianco, Devine, Hughes, & Banerjee, 2014).

Nevertheless, one problematic issue in the prevention of conduct problems has already been mentioned by Moffitt (1993). She determines that children who already show intraindividual vulnerability to conduct disorder from an early age on – such as in the form of neuropsychological deficits – are in great need of support and adequate parenting strategies by their caregivers. However, these children "evoke a challenge to even the most resourceful, loving, and patient families" (Moffitt, 1993, p. 9), and they evoke more negative reactions, for instance from their mothers, than normally developing children (Anderson, Lytton, & Romney, 1986). Furthermore, Moffitt (1993) assumes that because of the genetic heritability of vulnerabilities, children who are most in need of remedial cognitive stimulation often have parents who may not be able to provide it. Consequently, the children who are in the greatest need of supportive care do not receive it in many cases. Perhaps one solution could be to integrate programs that foster cognitive abilities like EF and ToM into classroom curricula, or even earlier into preschool programs. Naturally, this would probably not be enough to buffer the adverse effects of all intraindividual vulnerabilities and environmental variables, like coercive parenting. However, it might be one step forward in a series of different steps in the direction of preventing conduct problems and disorder in children. And, since ToM and EF are still developing in middle childhood and adolescence (e.g. Blakemore & Choudhury, 2006), fostering them would be a window of opportunity to prevent conduct problems.

Discussion

The prevention of conduct problems and fostering of self-control should be a target aim of public policies (Moffitt et al., 2011) to prevent the adverse long-term consequences associated with conduct problems and disorders in children, which are not only problematic for the individual child and his/her family, but are also a financial and social burden on the societal level (e.g., Farrington, 1995).

This thesis adds to the understanding of the role of specific cognitive deficits, ToM and EF, as risk factors for conduct problems, including different forms and functions of aggressive behavior in children. Thus, its results should be integrated into intervention programs aimed at preventing conduct problems in children.

8. Conclusion

In this thesis, deficits in ToM and EF were examined in tandem and separately as risk factors for conduct problems, including different forms and functions of aggressive behavior. The SIP theory by Crick and Dodge (1994) served as the theoretical background, besides other accounts like the amygdala theory of violent behavior (e.g., Blair et al., 2014), the developmental taxonomy of antisocial behavior (e.g. Moffitt, 1993), the problem-solving account (Séguin & Zelazo, 2005), and the frontal lobe hypothesis (Séguin, 2009). Further, a transactional perspective was adopted, which is suggested by Crick and Dodge (1994) and the ontogenic processes model of externalizing behavior (Beauchaine et al., 2015). All three reported studies and the additional analyses were based on a large community sample of $N = 1{,}657$ children, including three waves of a longitudinal study covering middle childhood and the transition to early adolescence (range 6 to 13 years) over a total of about three years. All data were analyzed with structural equation modeling. For ToM, a cognitive and an affective facet of ToM and, for EF, behavioral measures of working memory, inhibition, and set-shifting were assessed. Planning and emotional and inhibitory control, as well as conduct problems and forms and functions of aggressive behavior, were assessed through parent- and teacher-rated questionnaires.

The first research question was whether deficits in ToM are a risk factor for conduct problems, including different forms of aggressive behavior, from middle childhood through early adolescence. The results of Study 1 confirm the role of deficits in ToM as a longitudinal risk factor for teacher-rated physically and relationally aggressive behavior from a mean age of 8 to 11 years under the control of stable between-person differences in aggression. The results are in line with social information-processing theory (Crick & Dodge, 1994) and previous research (e.g., Kokkinos et al., 2016). In addition, we found a negative path from relationally aggressive behavior to later ToM, which confirms theoretical propositions on transactional relations by Crick and Dodge (1994) and the ontogenic processes model of externalizing behavior (Beauchaine et al., 2015). The results extend previous research in different ways: Past studies on the association between ToM and aggressive behavior were mostly cross-sectional, situated in preschool age and did not separate between distinct forms of aggressive behavior. Therefore, we examined the reciprocal relations between ToM and physically and relationally aggressive behavior in our longitudinal study that was situated in middle childhood over a period of three years. Further, we controlled for stable between-person differences in aggression and for past levels of ToM, which fosters a causal interpretation of the investigated paths (e.g., Hamaker et al., 2015). Second, we distinguished between two forms of aggressive behavior that have been found to differ between girls and boys in previous research (e.g. Card et al., 2008), and investigated gender differences in our paths. Finally, as one of the first studies, we also examined the reverse path from earlier aggressive behavior to later ToM.

The results of Study 3 confirm the role of deficits in ToM as a risk factor for parent-rated conduct problems cross-sectionally in an age range from 9 to 13 years. This also extends previous research, because most previous evidence on this association comes from early childhood (e.g., Hughes et al., 1998), and only little evidence comes from studies in middle childhood (e.g., Sharp, 2008). Until now, findings have been mixed, with some studies confirming a negative link between ToM and conduct problems, and others obtaining results that were difficult to interpret (Happé & Frith, 1996). As a consequence, Happé and Frith (1996) argued for a ‘theory of nasty minds,’ which means that children who show higher levels of conduct problems would use their ToM abilities in antisocial ways. The results of our study do not confirm this proposition, but rather indicate that children who have deficits in their ability to infer others’ cognitive and affective mental states tend to show higher levels of conduct problems. This finding is in line with social information-processing theory (Crick

& Dodge, 1994) and the amygdala theory of violent behavior (e.g., Blair et al., 2014). The latter states that children with deficits in their ToM might also show related deficits in empathy, which hinder an important mechanism inhibiting violent behavior. It is further proposed that children need to be able to infer the perspective of others to behave in moral ways (e.g., Hare, 1970), and consequently deficits in ToM could foster aggressive behavior in children. Furthermore, in both studies (Study 1 and 3) it was tested in separate multi-group models whether gender or three equally sized age groups moderated any of the examined paths. However, the underlying effects seem to be similar for females and males, as well as for the three age groups. To sum up the results for the first research question, deficits in ToM were found to be a risk factor for conduct problems, including different forms of aggressive behavior, longitudinally over a course of three years in middle childhood and cross-sectionally at the transition to adolescence.

The second research question was whether deficits in EF were a risk factor for conduct problems, including different forms and functions of aggressive behavior, from middle childhood through early adolescence. The results of Study 2 confirm that deficits in cool EF were a longitudinal risk factor for teacher-rated physically, relationally, and reactively aggressive behavior over a period of three years from a mean age of 8 to 11 years under the control of stable between-person differences in aggression. Deficits in EF were not related to proactively aggressive behavior. The path from cool EF to physical aggression, and as a trend also to relational aggression, was mediated by parent-rated habitual anger. Our findings replicate previous cross-sectional and longitudinal research from other age groups (e.g., Hughes & Ensor, 2008), and extend mixed cross-sectional results from middle childhood. Particularly previous evidence on the association between EF and relational aggression is mixed, with some studies finding a negative relation (e.g., McQuade et al., 2013), others a positive one (Poland et al., 2016), and some none (Diamantopoulou et al., 2007). However, our finding of a negative path is in line with the social information-processing theory of aggressive behavior (Crick & Dodge, 1994) that proposes that children who show physically as well as relationally aggressive behavior display deficits in their cognitive processing of social situations. Further, the negative paths from EF to physically and relationally aggressive behavior are in line with the problem-solving account (Séguin & Zelazo, 2005) and the frontal-lobe hypothesis (Séguin, 2009). The path from deficits in EF to later reactively aggressive behavior, and the nonsignificant path to later proactively aggressive behavior also confirm the few cross-sectional studies on the link between EF and the functions of aggression in middle childhood (e.g., Ellis et al., 2009; White et al., 2013). The finding of a negative, indirect effect from EF at T1 via anger at T2 to physical aggression at T3, and to relational aggression at T3 as a trend, also extends previous cross-sectional research (e.g., de Castro et al., 2005). This emphasizes the importance of considering emotions in the social-cognitive processing of children who show aggressive behavior (e.g., Lemerise & Arsenio, 2000). Our study was the first to uncover this mediational effect over a span of three years. In addition, all the investigated paths in our study held equally for females and males, as was tested in multi-group analyses.

Furthermore, the results of Study 3 indicate that deficits in cool EF had an indirect effect through deficits in ToM on conduct problems cross-sectionally in an age range of 9 to 13 years, but no direct effect of EF was found. This finding is in line with some previous research (e.g., Hughes & Ensor, 1998), but contradicts other research that also found a direct link between deficits in EF and conduct problems (e.g., Schoemaker et al., 2013). Different reasons could be responsible for the absence of a direct link between EF and conduct problems in our study, for instance, differences in participants' age, clinical status, or different measures. However, earlier reviews (Pennington & Ozonoff, 1996; Morgan & Lilienfeld, 2000) argued that children who show conduct disorder or problems could display deficits in neuropsychological tasks in general, and so deficits in EF could be a possible consequence, but not a sufficient condition for conduct problems. Nevertheless, the moderating roles of gender and three age groups of equal size were explored in multi-group analyses also for these paths, but as a result the effects remained similar for females and males despite differences in frequencies in some of the study variables, as well as for the three age groups.

Furthermore, following propositions of a developmental taxonomy of conduct disorder and problems (e.g. Moffitt, 1993), the role of deficits in cool and hot EF as a risk factor for a baseline level and the developmental trajectory of conduct problems was examined in the Additional Analyses. In latent growth curve analyses, a low and decreasing trajectory of conduct problems was determined from a mean age of 8 to 11 years. This result was in contrast to the assumption by Moffitt (e.g., 1993) that the prevalence rates of conduct disorder and problems would increase with the beginning of adolescence. One reason for that could be that at our last time point at a mean age of 11 years, for some children adolescence, and particularly puberty, might not have started yet. Moffitt et al. (2001) determined that the increase in conduct problems that is associated with adolescence-limited conduct problems is related to the timing of puberty in each individual. However, we could not test whether this was the reason for the decreasing trajectory of conduct problems in our study, or whether the floor effects in conduct problems were also responsible for it. We also tested the prediction of the baseline level and the trajectory of conduct problems by cool and hot EF. As a result, deficits in emotional and inhibitory control and planning were related to parent-rated conduct problems under the control of interindividual differences in conduct problems at a mean age of 8 years but did not predict the trajectory of conduct problems. All the variables were controlled for information-processing capacity, gender and age. These findings confirmed previous evidence on the relation between deficits in effortful control and conduct problems in middle childhood (e.g., Eisenberg et al., 2010). The results are also in line with the adapted social information-processing theory by Lemerise and Arsenio (2000), who proposed that emotion regulation processes mediate between social-cognitive processes and aggressive behavior, and the problem-solving account (Séguin & Zelazo, 2005). However, behaviorally measured cool EF predicted neither the baseline level nor the development of conduct problems in this study. Different reasons could be responsible for the missing link between behaviorally measured cool EF and conduct problems, such as a missing overlap in the source of the data, in contrast to conduct problems and inhibitory and emotional control that were all rated by parents. Also, other moderating variables like parenting practices (Eisenberg et al., 2015) that were not measured in the study, could be responsible for the missing link. This has to be determined in future studies.

In sum, the results for the second research question indicate that deficits in cool EF are a risk factor for later physical, relational, and reactive aggression, but not for proactive aggression over a course of three years in middle childhood. Further, deficits in EF were only linked indirectly through deficits in ToM to conduct problems cross-sectionally at the transition to early adolescence. Finally, deficits in emotional and inhibitory control and deficits in planning were related to the individual level of conduct problems at a mean age of eight years under the control of interindividual differences, but not deficits in behaviorally measured cool EF. Neither of them predicted the trajectory of conduct problems.

The third research question was whether deficits in ToM and in EF are risk factors for conduct problems from middle childhood through early adolescence, when examined in tandem. This was investigated in Study 3, and as a result, deficits in ToM were a risk factor when studied in tandem with deficits in cool EF, but deficits in cool EF were only indirectly through deficits in ToM related to conduct problems. This result is in line with the study by Hughes and Ensor (1998) and extends previous research, because deficits in ToM and EF have not often been studied in tandem and in one analysis. Additional multi-group analyses explored potential moderation of the investigated paths by gender and age group, however there was no indication for that. Consequently, the underlying effects seem to be similar for females and males, and for the three equally sized age groups.

Altogether, the results of all the conducted studies presented in this thesis extend previous research and confirm the propositions of the SIP model (e.g., Crick & Dodge, 1994), the amygdala theory of violent behavior (e.g., Blair et al., 2014), and partially the problem-solving account (Séguin & Zelazo, 2005) and the frontal-lobe hypothesis (Séguin, 2009). The results of the thesis suggest that deficits in ToM are a risk factor for conduct problems, including relational and physical aggression, and deficits in cool EF are a risk factor for physical, relational, and reactive aggression from middle childhood to early adolescence. Deficits in emotional and inhibitory control and planning seem to have a direct effect on conduct problems at a mean age of 8 years. When studied in tandem, deficits in cool EF seem to play an indirect role as a risk factor through deficits in ToM on conduct problems during the transition from middle childhood to early adolescence. This suggests that interventions preventing conduct problems should foster the ability to infer others' higher cognitive and affective mental states, as well as the ability to update information in working memory, to inhibit prepotent responses, to shift flexibly between different rule-sets, to plan behavior, and to emotionally and behaviorally control oneself in middle childhood through early adolescence. Interventions should address girls and boys equally and include the whole age range from 6 to 13 years, because the underlying effects seem to be similar for them. The reciprocal effects between deficits in ToM and relational aggression demonstrate the detrimental effect conduct problems can have on children's lives and health. The results of the thesis emphasize the need to intervene in the transactional processes between deficits in ToM and in EF and conduct problems, including different forms and functions of aggression, not only in early childhood, but also from middle and late childhood to early adolescence.

9. References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Anastassiou-Hadjicharalambous, X., & Warden, D. (2008). Cognitive and affective perspective-taking in conduct-disordered children high and low on callous-unemotional traits. *Child and Adolescent Psychiatry*, *2*(16), 1–11. http://doi.org/10.1186/1753-2000-2-16

Anderson, K. E., Lytton, H., & Romney, D. M. (1986). Mothers' interactions with normal and conduct-disordered boys: Who affects whom? *Developmental Psychology*, *22*, 604–609. http://doi.org/10.1037/0012-1649.22.5.604

Antonini, T. N., Becker, S. P., Tamm, L., & Epstein, J. N. (2015). Hot and cool executive functions in children with attention-deficit/hyperactivity disorder and comorbid oppositional defiant disorder. *Journal of the International Neuropsychological Society*, *21*, 584–595. http://doi.org/10.1017/S1355617715000752

Apperly, I. A., Warren, F., Andrews, B. J., Grant, J., & Todd, S. (2011). Developmental continuity in theory of mind: speed and accuracy of belief-desire reasoning in children and adults. *Child Development*, *82*, 1691–1703. http://doi.org/10.1111/j.1467-8624.2011.01635.x

Archer, J., & Coyne, S. M. (2005). An integrated review of indirect, relational, and social aggression. *Personality and Social Psychology Review*, *9*, 212–230. http://doi.org/10.1207/s15327957

Archibald, S. J., & Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychology*, *5*, 115–129. http://doi.org/10.1076/chin.5.2.115.3167

Arsenio, W. F., Cooperman, S., & Lover, A. (2000). Affective predictors of preschoolers' aggression and peer acceptance: Direct and indirect effects. *Developmental Psychology*, *36*, 438–448. http://doi.org/10.1037//0012-1649.36.4.438

Arsenio, W. F., & Fleiss, K. (1996). Typical and behaviourally disruptive children's understanding of the emotional consequences of socio-moral events. *British Journal of Developmental Psychology*, *14*, 173–186. http://doi.org/10.1111/j.2044-835X.1996.tb00700.x

August, G. J., Piehler, T. F., & Bloomquist, M. L. (2016). Being "SMART" about adolescent conduct problems prevention: Executing a SMART pilot study in a juvenile diversion agency. *Journal of Clinical Child & Adolescent Psychology*, *45*, 495–509. http://doi.org/10.1080/15374416.2014.945212

Austin, G., Bondü, R., & Elsner, B. (2017). Longitudinal relations between children's cognitive and affective theory of mind with reactive and proactive aggression. *Aggressive Behavior*, *43*, 440–449. http://doi.org/10.1002/ab.21702

Austin, G., Groppe, K., & Elsner, B. (2014). The reciprocal relationship between executive function and theory of mind in middle childhood: a 1-year longitudinal perspective. *Frontiers in Psychology*, *5*(655). http://doi.org/10.3389/fpsyg.2014.00655

Ayduk, O., Rodriguez, M. L., Mischel, W., Shoda, Y., & Wright, J. (2007). Verbal intelligence and self-regulatory competencies: Joint predictors of boys' aggression. *Journal of Research in Personality*, *41*, 374–388. http://doi.org/10.1016/j.jrp.2006.04.008

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47–89. http://doi.org/https://doi.org/10.1016/S0079-7421(08)60452-1

Barker, E. D., & Maughan, B. (2009). Differentiating early-onset persistent versus childhood-limited conduct problem youth. *American Journal of Psychiatry*, *166*, 900–908. http://doi.org/10.1176/appi.ajp.2009.08121770

Baron-Cohen. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences*, *6*(6), 248–254. http://doi.org/10.1016/S1364-6613(02)01904-6

Baron-Cohen, S., Ring, H. A., Ring, H. A., Wheelright, S., … Williams, S. C. (1999). Social intelligence in the normal and autisitic brain: an fMRI study. *European Journal of Neuroscience*, *11*, 1891–1898. http://doi.org/10.1046/j.1460-9568.1999.00621.x

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). A new test of social sensitivity: Detection of faux pas in normal children and children with Asperger syndrome. *Journal of Autism and Developmental Disorders*, *29*, 407–418.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *42*, 241–251. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11280420

Baron, R. A., & Richardson, D. R. (1994). *Human aggression*. New York, NY: Plenum.

Beauchaine, T. P., & Hinshaw, S. P. (2015). Preface. In T. P. Beauchaine & S. P. Hinshaw (Eds.), *The Oxford handbook of externalizing spectrum disorders* (pp. 1–7). New York, NY: Oxford University. http://doi.org/10.1093/oxfordhb/9780199219322.003.0023

Beauchaine, T. P., Hong, J., & Marsh, P. (2008). Sex differences in autonomic correlates of conduct problems and aggression. *Journal of the American Academy of Child and Adolescent Psychiatry*, *47*, 788–796. http://doi.org/10.1097/CHI.Ob013e318172ef4b

Beauchaine, T. P., & McNulty, T. (2013). Comorbidities and continuities as ontogenic processes: Toward a developmental spectrum model of externalizing psychopathology. *Development and Psychopathology*, *25*, 1505–1528. http://doi.org/10.1017/S0954579413000746

Beauchaine, T. P., Shader, T. M., & Hinshaw, S. P. (2015). An ontogenic processes model of externalizing psychopathology. In T. P. Beauchaine & S. P. Hinshaw (Eds.), *The Oxford handbook of externalizing spectrum disorders* (pp. 485–502). New York, NY: Oxford University. http://doi.org/10.1093/oxfordhb/9780199324675.013.15

Becker, M. G., Isaac, W., & Hynd, G. W. (1987). Neuropsychological development of nonverbal behaviors attributed to "frontal lobe" functioning. *Developmental Neuropsychology*, *3*, 275–

298. http://doi.org/10.1080/87565648709540381

Berkowitz, L. (1989). Frustration-aggression hypothesis: Examination and reformulation. *Psychological Bulletin*, *106*(1), 59–73. http://doi.org/10.1037/0033-2909.106.1.59

Berkowitz, L. (2012). A different view of anger: The cognitive-neoassociation conception of the relation of anger to aggression. *Aggressive Behavior*, *38*, 322–333. http://doi.org/10.1002/ab.21432

Berkowitz, L., & Harmon-Jones, E. (2004). Toward an understanding of the determinants of anger. *Emotion*, *4*, 107–130. http://doi.org/10.1037/1528-3542.4.2.107

Berlin, L., & Bohlin, G. (2002). Response inhibition, hyperactivity, and conduct problems among preschool children. *Journal of Clinical Child and Adolescent Psychology*, *31*, 242–251. http://doi.org/10.1207/S15374424JCCP3102_09

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, *81*, 1641–1660. http://doi.org/10.1111/j.1467-8624.2010.01499.x

Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review*, *29*, 180–200. http://doi.org/10.1016/j.dr.2009.05.002

Bevilacqua, L., Hale, D., Barker, E. D., & Viner, R. (2017). Conduct problems trajectories and psychosocial outcomes: a systematic review and meta-analysis. *European Child & Adolescent Psychiatry*, *advance online*, 1–22. http://doi.org/10.1007/s00787-017-1053-4

Björkqvist, K., Österman, K., & Kaukiainen, A. (1992). The development of direct and indirect aggressive strategies in males and females. In K. Björkqvist & P. Niemelä (Eds.), *Of mice and women: Aspects of female aggression.* (pp. 51–64). London, UK: Academic Press.

Blair, R. J. R. (2005). Applying a cognitive neuroscience perspective to the disorder of psychopathy. *Development and Psychopathology*, *17*, 865–891. http://doi.org/10.1017/S0954579405050418

Blair, R. J. R., Leibenluft, E., & Pine, D. S. (2014). Conduct disorder and callous-unemotional traits in youth. *The New England Journal of Medicine*, *371*, 2207–2216.

http://doi.org/10.1056/NEJMra1315612

Blakemore, S.-J., & Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *47*, 296–312. http://doi.org/10.1111/j.1469-7610.2006.01611.x

Bock, A. M., Gallaway, K. C., & Hund, A. M. (2014). Specifying links between executive functioning and theory of mind during middle childhood: Cognitive flexibility predicts social understanding. *Journal of Cognition and Development*, *16*, 509–521. http://doi.org/10.1080/15248372.2014.888350

Bosacki, S., & Astington, J. W. (1999). Theory of mind in preadolescence: Relations between social understanding and social competence. *Social Development*, *8*, 237–255. http://doi.org/10.1111/1467-9507.00093

Brendgen, M., Vitaro, R., Tremblay, R. E., Lavoie, F., Vitaro, F., Tremblay, R. E., & Lavoie, F. (2001). Reactive and proactive aggression: Predictions to physical violence in different contexts and moderating effects of parental monitoring and caregiving behavior. *Journal of Abnormal Child Psychology*, *29*, 293–304. http://doi.org/10.1023/A:1010305828208

Bridgett, D. J., Oddi, K. B., Laake, L. M., Murdock, K. W., & Bachmann, M. N. (2013). Integrating and differentiating aspects of self-regulation: Effortful control, executive functioning, and links to negative affectivity. *Emotion*, *13*, 47–63. http://doi.org/10.1037/a0029536

Brocki, K. C., & Bohlin, G. (2004). Executive functions in children aged 6 to 13: a dimensional and developmental study. *Developmental Neuropsychology*, *26*, 571–593. http://doi.org/10.1207/s15326942dn2602_3

Burke, J. D., Loeber, R., Lahey, B. B., & Rathouz, P. J. (2005). Developmental transitions among affective and behavioral disorders in adolescent boys. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *46*, 1200–1210. http://doi.org/10.1111/j.1469-7610.2005.00422.x

References

Calero, C. I., Salles, A., Semelman, M., & Sigman, M. (2013). Age and gender dependent development of theory of mind in 6- to 8-years old children. *Frontiers in Human Neuroscience*, *7*(281). http://doi.org/10.3389/fnhum.2013.00281

Cantin, R. H., Gnaedinger, E. K., Gallaway, K. C., Hesson-McInnis, M. S., & Hund, A. M. (2016). Executive functioning predicts reading, mathematics, and theory of mind during the elementary years. *Journal of Experimental Child Psychology*, *146*, 66–78. http://doi.org/10.1016/j.jecp.2016.01.014

Capage, L., & Watson, A. C. (2001). Individual differences in theory of mind, aggressive behavior, and social skills in young children. *Early Education & Development*, *12*, 613–628.

Card, N. A., & Little, T. D. (2006). Proactive and reactive aggression in childhood and adolescence: A meta-analysis of differential relations with psychosocial adjustment. *International Journal of Behavioral Development*, *30*, 466–480. http://doi.org/10.1177/0165025406071904

Card, N. A., Stucky, B. D., Sawalani, G. M., & Little, T. D. (2008). Direct and indirect aggression during childhood and adolescence: a meta-analytic review of gender differences, intercorrelations, and relations to maladjustment. *Child Development*, *79*, 1185–1229. http://doi.org/10.1111/j.1467-8624.2008.01184.x

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, *28*, 595–616. http://doi.org/10.1207/s15326942dn2802_3

Carlson, S. M., Claxton, L. J., & Moses, L. J. (2015). The relation between executive function and theory of mind is more than skin deep. *Journal of Cognition and Development*, *16*, 186–197. http://doi.org/10.1080/15248372.2013.824883

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*, 1032–1053. http://doi.org/10.1111/1467-

8624.00333

Carlson, S. M., & Wang, T. S. (2007). Inhibitory control and emotion regulation in preschool children. *Cognitive Development*, *22*, 489–510. http://doi.org/10.1016/j.cogdev.2007.08.002

Caspi, A. (1996). Behavioral observations at age 3 years predict adult psychiatric disorders. *Archives of General Psychiatry*, *53*, 1033–1039. http://doi.org/10.1001/archpsyc.1996.01830110071009

Chapman, E., Baron-Cohen, S., Auyeung, B., Knickmeyer, R., Taylor, K., & Hackett, G. (2006). Fetal testosterone and empathy: Evidence from the empathy quotient (EQ) and the "Reading the Mind in the Eyes" Test. *Social Neuroscience*, *1*, 135–148. http://doi.org/10.1080/17470910600992239

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*, 2245–2256. http://doi.org/10.1037/a0032293

Choudhury, S., Blakemore, S.-J., & Charman, T. (2006). Social cognitive development during adolescence. *Social Cognitive and Affective Neuroscience*, *1*(3), 165–174. http://doi.org/10.1093/scan/nsl024

Cicchetti, D., & Rogosch, F. A. (1996). Equifinality and multifinality in developmental psychopathology. *Development and Psychopathology*, *8*, 597-600. http://doi.org/10.1017/S0954579400007318

Collins, W. A. (1984). *Development during middle childhood - The years from six to twelve*. (W. A. Collins, Ed.). Washington, DC: National Academy Press.

Collishaw, S., Maughan, B., Goodman, R., & Pickles, A. (2004). Time trends in adolescent mental health. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *45*, 1350–1362. http://doi.org/10.1111/j.1469-7610.2004.00335.x

Conklin, H. M., Luciana, M., Hooper, C. J., & Yarger, R. S. (2007). Working memory performance in typically developing children and adolescents: Behavioral evidence of protracted frontal lobe development. *Developmental Neuropsychology*, *31*, 103–128. http://doi.org/10.1207/s15326942dn3101_6

Côté, S. M., Vaillancourt, T., Barker, E. D., Nagin, D., & Tremblay, R. E. (2007). The joint development of physical and indirect aggression: Predictors of continuity and change during childhood. *Development and Psychopathology*, *19*, 37–55. http://doi.org/10.1017/S0954579407070034

Crean, H. F., & Johnson, D. B. (2013). Promoting alternative thinking strategies (PATHS) and elementary school aged children's aggression: Results from a cluster randomized trial. *American Journal of Community Psychology*, *52*, 56–72. http://doi.org/10.1007/s10464-013-9576-4

Crick, N. R. (1996). The role of overt aggression, relational aggression, and prosocial behavior in the prediction of children's future social adjustment. *Child Development*, *67*, 2317–2327. http://doi.org/10.1111/j.1467-8624.1996.tb01859

Crick, N. R., Casas, J. F., & Ku, H.-C. (1999). Relational and physical forms of peer victimization in preschool. *Developmental Psychology*, *35*, 376–385. http://doi.org/10.1037/0012-1649.35.2.376

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 mechanisms in reactive and proactive aggression. *Child Development*, *67*(3), 993–1002. http://doi.org/http://dx.doi.org/10.2307/1131875

Crick, N. R., & Grotpeter, J. K. (1995). Relational aggression, gender, and social-psychological adjustment. *Child Development*, *66*, 710–722. http://doi.org/10.2307/1131945

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*, 1134–1142.

Crick, N. R., Werner, N. E., Casas, J. F., O´Brien, K. M., Nelson, D. A., Grotpeter, J. K., & Markon, K. (1999). Childhood aggression and gender: A new look at an old problem. In D. Bernstein (Ed.), *Nebraska symposium on motivation* (pp. 75–141). Lincoln, NE: University of Nebraska Press.

Cui, L., Colasante, T., Malti, T., Ribeaud, D., & Eisner, M. P. (2016). Dual trajectories of reactive and proactive aggression from mid-childhood to early adolescence: Relations to sensation seeking, risk taking, and moral reasoning. *Journal of Abnormal Child Psychology*, *44*, 663–675. http://doi.org/10.1007/s10802-015-0079-7

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*, 105–116. http://doi.org/10.1207/s15374424jccp3401

de Castro, B. O., & van Dijk, A. (2018). "It's gonna end up with a fight anyway". Social cognitive processes in children with disruptive behavior disorders. In J. E. Lochman & W. Matthys (Eds.), *The Wiley Handbook of disruptive and impulse-control disorders.* (pp. 237–253). Hoboken, NJ: Wiley.

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*, 916–934. http://doi.org/10.1111/1467-8624.00447

de la Osa, N., Granero, R., Domenech, J. M., Shamay-Tsoory, S., & Ezpeleta, L. (2016). Cognitive and affective components of theory of mind in preschoolers with oppositional defiance disorder: Clinical evidence. *Psychiatry Research*, *241*, 128–134. http://doi.org/10.1016/j.psychres.2016.04.082

Devine, R. T., & Hughes, C. (2013). Silent films and strange stories: Theory of mind, gender, and social experiences in middle childhood. *Child Development*, *84*, 989–1003.

http://doi.org/10.1111/cdev.12017

Devine, R. T., & Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: A meta-analysis. *Child Development*, *85*, 1777–1794. http://doi.org/10.1111/cdev.12237

Devine, R. T., White, N., Ensor, R., & Hughes, C. (2016). Theory of mind in middle childhood: Longitudinal associations with executive function and social competence. *Developmental Psychology*, *52*, 758–771. http://doi.org/10.1037/dev0000105

Diamantopoulou, S., Rydell, A. M., Thorell, L. B., & Bohlin, G. (2007). Impact of executive functioning and symptoms of attention deficit hyperactivity disorder on children’s peer relations and school performance. *Developmental Neuropsychology*, *32*, 521–542. http://doi.org/10.1080/87565640701360981

Diamond, A. (2013). Executive functions. *The Annual Review of Psychology*, *64*(9), 135–168. http://doi.org/10.1146/annurev-psych-113011-143750

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, *333*(6045), 959–964. http://doi.org/10.1126/science.1204529

Dodge, K. A. (2006). Translational science in action: Hostile attributional style and the development of aggressive behavior problems. *Development and Psychopathology*, *18*, 791–814. http://doi.org/10.10170S0954579406060391

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

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*, 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, C. (2015). Hostile attributional bias and aggressive behavior in global context. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 9310–9315. http://doi.org/10.1073/pnas.1418572112

Dodge, K. A., Pettit, G. S., & Bates, J. E. (1994). Socialization mediators of the relation between socioeconomic status and child conduct problems. *Child Development*, *65*, 649–665.

Dolan, M., & Lennox, C. (2013). Cool and hot executive function in conduct-disordered adolescents with and without co-morbid attention deficit hyperactivity disorder: relationships with externalizing behaviours. *Psychological Medicine*, *43*, 2427–2436. http://doi.org/10.1017/S0033291712003078

Drabick, D., Steinberg, E., & Hampton, A. (2015). Overview of DSM disruptive behavior disorders. In T. P. Beauchaine & S. P. Hinshaw (Eds.), *The Oxford Handbook of Externalizing Spectrum Disorders* (pp. 3–18). New York, NY: Oxford University. http://doi.org/10.1093/oxfordhb/9780199324675.013.16

Dumontheil, I., Apperly, I. A., & Blakemore, S.-J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science*, *13*, 331–338. http://doi.org/10.1111/j.1467-7687.2009.00888.x

Edelbrock, C., Rende, R., Plomin, R., & Thompson, L. A. (1995). A twin study of competence and problem behavior in childhood and early adolescence. *Journal of Child Psychology and Psychiatry*, *36*, 775–785. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7559844

Eisenberg, N., Fabes, R. A., Murphy, B. C., Shepard, S., Guthrie, I. K., Mazsk, P., … Jones, S. (1999). Prediction of elementary school children's socially appropriate and problem behavior from anger reactions at age 4-6 years. *Journal of Applied Developmental Psychology*, *20*, 119–142. http://doi.org/10.1016/S0193-3973(99)80007-0

References

Eisenberg, N., & Spinrad, T. L. (2004). Emotion-related regulation: Sharpening the definition. *Child Development*, *75*, 334–339.

Eisenberg, N., Spinrad, T. L., & Eggum, N. D. (2010). Emotion-related self-regulation and its relation to children's maladjustment. *Annual Review of Clinical Psychology*, *6*, 495–525. http://doi.org/10.1146/annurev.clinpsy.121208.131208

Eisenberg, N., Spinrad, T. L., Fabes, R. A., Reiser, M., Cumberland, A., Stephanie, A., … Thompson, M. (2004). The relations of effortful control and impulsivity to children's resiliency and adjustment. *Child Development*, *75*, 25–46.

Eisenberg, N., Taylor, Z. E., Widaman, K. F., & Spinrad, T. L. (2015). Externalizing symptoms, effortful control, and intrusive parenting: A test of bidirectional longitudinal relations during early childhood. *Development and Psychopathology*, *27*, 953–968. http://doi.org/10.1017/S0954579415000620

Elliot, D. S., & Menard, S. (1996). Delinquent friends and delinquent behavior: Temporal and developmental patterns. In J. D. Hawkings (Ed.), *Cambridge criminology series. Delinquency and crime: Current Theories.* (pp. 28–67). New York, NY: Cambridge University.

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*, 945–956. http://doi.org/10.1007/s10802-009-9321-5

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Press.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*, 430–457. http://doi.org/10.1207/s15328007sem0803_5

Epps, E. G., & Smith, S. F. (1984). School and children: The middle childhood years. In W. A. Collins (Ed.), *Development during middle childhood: The years from six to twelve* (pp.

283–334). Washington, DC: National Academy Press.

Fahie, C. M., & Symons, D. K. (2003). Executive functioning and theory of mind in children clinically referred for attention and behavior problems. *Journal of Applied Developmental Psychology*, *24*, 51–73. http://doi.org/10.1016/S0193-3973(03)00024-8

Fairchild, G., Van Goozen, S. H. M., Calder, A. J., & Goodyer, I. M. (2013). Research review: Evaluating and reformulating the developmental taxonomic theory of antisocial behaviour. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *54*, 924–940. http://doi.org/10.1111/jcpp.12102

Farrington, D. P. (1995). The development of offending and antisocial behaviour from childhood: Key findings from the Cambridge study in delinquent development. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *36*, 929–964. http://doi.org/10.1111/j.1469-7610.1995.tb01342.x

Fite, P. J., Colder, C. R., Lochman, J. E., & Wells, K. C. (2008). Developmental trajectories of proactive and reactive aggression from fifth to ninth grade. *Journal of Clinical Child and Adolescent Psychology*, *37*, 412–421. http://doi.org/10.1080/15374410801955920

Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect. *Merrill-Palmer Quarterly*, *50*, 274–290.

Fontaine, N. M. G., Rijsdijk, F. V., McCrory, E. J. P., & Viding, E. (2010). Etiology of different developmental trajectories of callous-unemotional traits. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*, 656–664. http://doi.org/10.1016/j.jaac.2010.03.014

Frick, P. J., & Morris, A. S. (2004). Temperament and developmental pathways to conduct problems. *Journal of Clinical Child and Adolescent Psychology*, *33*(1), 21–31. http://doi.org/10.1207/S15374424JCCP3301

Frith, C. D., & Frith, U. (2012). Mechanisms of social cognition. *Annual Review of Psychology*, *63*, 287–313. http://doi.org/10.1146/annurev-psych-120710-100449

Frith, U., & Frith, C. (2010). The social brain: allowing humans to boldly go where no other species has been. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *365*(1537), 165–176. http://doi.org/10.1098/rstb.2009.0160

Frye, D., Zelazo, P. D., & Burack, J. (1998). Cognitive complexity and control: I. Theory of mind in typical and atypical development. *Current Directions in Psychological Science*, *7*, 116–121.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, *10*, 483–527. http://doi.org/https://doi.org/10.1016/0885-2014(95)90024-1

Gagne, J. R., & Hill Goldsmith, H. (2011). A longitudinal analysis of anger and inhibitory control in twins from 12 to 36 months of age. *Developmental Science*, *14*, 112–124. http://doi.org/10.1111/j.1467-7687.2010.00969.x

Gallagher, A. (1998). Gender and antecedents of performance in mathematics testing. *Teachers College Record*, *100*, 297–314.

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychological Bulletin*, *134*, 31–60. http://doi.org/10.1037/0033-2909.134.1.31

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*, 177–190. http://doi.org/10.1037/0012-1649.40.2.177

Gatzke-Kopp, L. M., Greenberg, M. T., Fortunato, C. K., & Coccia, M. A. (2012). Aggression as an equifinal outcome of distinct neurocognitive and neuroaffective processes. *Development and Psychopathology*, *24*, 985–1002. http://doi.org/10.1017/S0954579412000491

Geiser, G. (2011). *Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung* (2.). Wiesbaden: VS Verlag für Sozialwissenschaften.

Gini, G. (2006). Social cognition and moral cognition in bullying: what's wrong? *Aggressive Behavior*, *32*, 528–539. http://doi.org/10.1002/ab.20153

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior Rating of Executive*

Function. Lutz, FL: Psychological Assessment Resources.

Golan, O., Baron-Cohen, S., & Hill, J. (2006). The Cambridge mindreading (CAM) face-voice battery: Testing complex emotion recognition in adults with and without Asperger syndrome. *Journal of Autism and Developmental Disorders*, *36*, 169–183. http://doi.org/10.1007/s10803-005-0057-y

Goldstein, S., & Naglieri, J. A. (2014). *Handbook of Executive Functioning*. http://doi.org/10.1007/978-0-387-79948-3

Gomez-Garibello, C., & Talwar, V. (2015). Can you read my mind? Age as a moderator in the relationship between theory of mind and relational aggression. *International Journal of Behavioral Development*, *39*, 552–559. http://doi.org/10.1177/0165025415580805

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *38*, 581–586.

Granvald, V., & Marciszko, C. (2016). Relations between key executive functions and aggression in childhood. *Child Neuropsychology*, *22*(5), 537–555. http://doi.org/10.1080/09297049.2015.1018152

Grueneisen, S., Wyman, E., & Tomasello, M. (2015). "I know you don't know I know..." Children use second-order false-belief reasoning for peer coordination. *Child Development*, *86*, 287–293. http://doi.org/10.1111/cdev.12264

Hagekull, B., & Bohlin, G. (1998). Preschool temperament and environmental factors related to the five-factor model of personality in middle childhood. *Merrill-Palmer Quarterly*, *44*(2), 194–215.

Halpern, D. F., & Lamay, M. L. (2000). The smarter sex: A critical review of sex differences in intelligence. *Educational Psychology Review*, *12*, 229–246. http://doi.org/10.1023/A:1009027516424

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*, 102–116. http://doi.org/10.1037/a0038889

References

Hansen Lagattuta, K., Sayfan, L., & Harvey, C. (2014). Beliefs about thought probability: Evidence for persistent errors in mindreading and links to executive control. *Child Development*, *85*, 659–674. http://doi.org/10.1111/cdev.12154

Happé, F., & Frith, U. (1996). Theory of mind and social impairment in children with conduct disorder. *British Journal of Developmental Psychology*, *14*, 385–398. http://doi.org/10.1111/j.2044-835X.1996.tb00713.x

Happé, F. G. E. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, *24*, 129–154. http://doi.org/10.1007/BF02172093

Hare, R. D. (1970). *Psychopathy: Theory and research*. Oxford, UK: Wiley & Sons.

Hartup, W. W. (1984). The peer context in middle childhood. In W. A. Collins (Ed.), *Development during middle childhood: The years from six to twelve* (pp. 240–282). Washington, DC: National Academy Press.

Hawes, S. W., Perlman, S. B., Byrd, A. L., Raine, A., Loeber, R., & Pardini, D. A. (2016). Chronic anger as a precursor to adult antisocial personality features : The moderating influence of cognitive control. *Journal of Abnormal Child Psychology*, *125*, 64–74. http://doi.org/10.1037/abn0000129.

Healey, D. M., Marks, D. J., & Halperin, J. M. (2011). Examining the interplay among negative emotionality, cognitive functioning, and attention deficit/hyperactivity disorder symptom severity. *Journal of the International Neuropsychological Society*, *17*, 502–510. http://doi.org/10.1017/S1355617711000294

Henington, C., Hughes, J. N., Cavell, T. A., & Thompson, B. (1998). The role of relational aggression in identifying aggressive boys and girls. *Journal of School Psychology*, *36*, 457–477. http://doi.org/10.1016/S0022-4405(98)00015-6

Hoffman, M. L. (1987). The contribution of empathy to justice and moral judgment. In N.

Eisenberg & J. Strayer (Eds.), *Empathy and its development* (pp 45-68). Cambridge, NY: Cambridge University Press.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424–453. http://doi.org/10.1037/1082-989X.3.4.424

Huang-Pollock, C., Shapiro, Z., Galloway-Long, H., & Weigard, A. (2017). Is poor working memory a transdiagnostic risk factor for psychopathology? *Journal of Abnormal Child Psychology*, *45*, 1477–1490. http://doi.org/10.1007/s10802-016-0219-8

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., Romano, L. J., McAuliffe, M. D., & Morrow, M. T. (2010). Anger and the reactive-proactive aggression distinction in childhood and adolescence. In M. Potegal, G. Stemmler, & C. D. Spielberger (Eds.), *International handbook of anger: constituent and concomitant biological, psychological, and social processes* (pp. 231–239). New York: Springer.

Hubbard, J. A., Smithmyer, C. M., Ramsden, S. R., Elizabeth, H., Flanagan, K. D., Dearing, K. F., … Simons, R. F. (2002). Observational, physiological, and self-report measures of children's anger: Relations to reactive versus proactive aggression. *Child Development*, *73*, 1101–1118. http://doi.org/10.1111/1467-8624.00460

Huesmann, L. R., Eron, L. D., & Yarmel, P. W. (1987). Intellectual functioning and aggression. *Journal of Personality and Social Psychology*, *52*(1), 232–240. http://doi.org/10.1037/0022-3514.52.1.232

Huesmann, R. L. (1998). The role of social information processing and cognitive schema in the acquisition and maintenance of habitual aggressive behavior. In R. G. Geen & E.

Donnerstein (Eds.), *Human aggression. Theories, research, and implications for social policy* (pp. 73–109). San Diego, CA: Academic Press.

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology*, *16*, 233–253. http://doi.org/10.1111/j.2044-835X.1998.tb00921.x

Hughes, C., & Devine, R. T. (2015). Individual differences in theory of mind from preschool to adolescence: Achievements and directions. *Child Development Perspectives*, *9*, 149–153. http://doi.org/10.1111/cdep.12124

Hughes, C., Dunn, J., & White, A. (1998). Trick or treat? Uneven understanding of mind and emotion and executive dysfunction in “hard-to-manage” preschoolers. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *39*, 981–994. http://doi.org/9804031

Hughes, C., & Ensor, R. (2006). Behavioural problems in 2-year-olds: Links with individual differences in theory of mind, executive function and harsh parenting. *Journal of Child Psychology and Psychiatry*, *47*, 488–497. http://doi.org/10.1111/j.1469-7610.2005.01519.x

Hughes, C., & Ensor, R. (2008). Does executive function matter for preschoolers’ problem behaviors? *Journal of Abnormal Child Psychology*, *36*, 1–14. http://doi.org/10.1007/s10802-007-9107-6

Hughes, C., Ensor, R., & Marks, A. (2011). Individual differences in false belief understanding are stable from 3 to 6 years of age and predict children’s mental state talk with school friends. *Journal of Experimental Child Psychology*, *108*(1), 96–112. http://doi.org/https://doi.org/10.1016/j.jecp.2010.07.012

Hughes, C., Graham, A., & Grayson, A. (2005). Executive function in childhood: Development and disorder. In I. J. Oates (Ed.), *Cognitive development* (pp. 205–230). Milton Keynes, UK: Open University Press.

Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social Development*, *13*, 590–619. http://doi.org/10.1111/j.1467-9507.2004.00285.x

Hughes, C., White, A., Sharpen, J., & Dunn, J. (2000). Antisocial, angry, and unsympathetic: "hard-to-manage" preschoolers' peer problems and possible cognitive influences. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *41*, 169–179. http://doi.org/10.1111/1469-7610.00558

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*, 2017–2036. http://doi.org/10.1016/j.neuropsychologia.2006.01.010

Huyder, V., Nilsen, E. S., & Bacso, S. A. (2017). The relationship between children's executive functioning, theory of mind, and verbal skills with their own and others' behaviour in a cooperative context: Changes in relations from early to middle school-age. *Infant and Child Development*, *e2027*(advance online). http://doi.org/10.1002/icd.2027

Ibanez, A., Huepe, D., Gempp, R., Gutiérrez, V., Rivera-Rei, A., & Toledo, M. I. (2013). Empathy, sex and fluid intelligence as predictors of theory of mind. *Personality and Individual Differences*, *54*, 616–621. http://doi.org/10.1016/j.paid.2012.11.022

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, *17*, 213–233. http://doi.org/10.1007/s11065-007-9040-z

Keenan, K., Wroblewski, K., Hipwell, A., & Stouthamer-loeber, M. (2010). Age of onset, symptom threshold, and expansion of the nosology of conduct disorder for girls. *Journal of Abnormal Child Psychology*, *119*, 689–698. http://doi.org/10.1037/a0019346

Keulers, E. H. H., Evers, E. A. T., Stiers, P., & Jolles, J. (2010). Age, sex, and pubertal phase influence mentalizing about emotions and actions in adolescents. *Developmental Neuropsychology*, *35*(5), 555–569. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2010-17435-

007&site=ehost-live

Kimonis, E. R., & Frick, P. J. (2010). Etiology of oppositional defiant disorder and conduct disorder: Biological, familial and environmental factors identified in the development of disruptive behavior disorders. In R. C. Murrihy, A. D. Kidman, & T. H. Ollendick (Eds.), *Clinical handbook of assessing and treating conduct problems in youth* (pp. 49–76). New York, NY: Springer.

Klasen, H., Woerner, W., Wolke, D., Meyer, R., Overmeyer, S., & Kaschnitz, W. Rothenberger, A. Goodman, R. (2000). Comparing the German versions of the Strengths and Difficulties Questionnaire (SDQ-Deu) and the Child Behavior Checklist. *European Child and Adolescent Psychiatry*, *9*, 271–276.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.

Kokkinos, C. M., Voulgaridou, I., Mandrali, M., & Parousidou, C. (2016). Interactive links between relational aggression, theory of mind, and moral disengagement among early adolescents. *Psychology in the Schools*, *53*, 253–269. http://doi.org/10.1002/pits.21902

Kouklari, E.-C., Thompson, T., Monks, C. P., & Tsermentseli, S. (2017). Hot and cool executive function and its relation to theory of mind in children with and without autism spectrum disorder. *Journal of Cognition and Development*, *18*, 399–418. http://doi.org/10.1080/15248372.2017.1339708

Krahé, B. (2013). *The social psychology of aggression* (2nd ed.). New York, NY: Psychology Press.

Lahey, B. B., Schwab-Stone, M., Goodman, S. H., Waldman, I. D., Canino, G., Rathouz, P. J., … Jensen, P. S. (2000). Age and gender differences in oppositional behavior and conduct problems: A cross-sectional household study of middle childhood and adolescence. *Journal of Abnormal Psychology*, *109*, 488–503. http://doi.org/10.1037/0021-843X.109.3.488

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*, 715–724. 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*, 593–602. http://doi.org/10.1017/S0954579410000301

Lansford, J. E., Skinner, A. T., Sorbring, E., Giunta, L. Di, Deater-Deckard, K., Dodge, K. A., … Chang, L. (2012). Boys' and girls' relational and physical aggression in nine countries. *Aggressive Behavior*, *38*, 298–308. http://doi.org/10.1002/ab.21433

Lecce, S., Bianco, F., Devine, R. T., & Hughes, C. (2017). Relations between theory of mind and executive function in middle childhood: A short-term longitudinal study. *Journal of Experimental Child Psychology*, *163*, 69–86. http://doi.org/10.1016/j.jecp.2017.06.011

Lecce, S., Bianco, F., Devine, R. T., Hughes, C., & Banerjee, R. (2014). Promoting theory of mind during middle childhood: a training program. *Journal of Experimental Child Psychology*, *126*, 52–67. http://doi.org/10.1016/j.jecp.2014.03.002

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*, 59–80. http://doi.org/10.1348/026151003321164627

Leibenluft, E. (2011). Severe mood dysregulation, irritability, and the diagnostic boundaries of bipolar disorder in youths. *American Journal of Psychiatry*, *168*, 129–142. http://doi.org/10.1176/appi.ajp.2010.10050766

Leibenluft, E. (2017). Irritability in children: what we know and what we need to learn. *World Psychiatry*, *16*(1), 100–101. http://doi.org/10.1002/wps.20397

Leibenluft, E., & Stoddard, J. (2013). The developmental psychopathology of irritability. *Development and Psychopathology*, *25*, 1473–1487.

http://doi.org/10.1017/S0954579413000722

Lemerise, E. A., & Arsenio, W. F. (2000). An integrated model of emotion processes and cognition in social information processing. *Child Development*, *71*, 107–118. http://doi.org/10.1111/1467-8624.00124

Loeber, R., Burke, J. D., Lahey, B. B., Winters, A., & Zera, M. (2000). Oppositional defiant and conduct disorder: A review of the past 10 years, part I. *Journal of American Academic Child Adolescent Psychiatry*, *39*, 1468–1482.

Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005). The development of nonverbal working memory and executive control processes in adolescents. *Child Development*, *76*, 697–712. Retrieved from http://www.jstor.org/stable/3696460

Luciana, M., & Nelson, C. a. (1998). The functional emergence of prefrontally-guided working memory systems in four- to eight-year-old children. *Neuropsychologia*, *36*, 273–293.

Maccoby, E. E. (1990). Gender and relationships. A developmental account. *The American Psychologist*, *45*(4), 513–520. http://doi.org/10.1037/0003-066X.45.4.513

Markus, H. J., & Nurius, P. S. (1984). Self-understanding and self-regulation in middle childhood. In W. A. Collins (Ed.), *Development during middle childhood: The years from six to twelve* (pp. 147–183). Washington, DC: National Academy Press.

Marmor, G. S., & Montemayor, R. (1977). The cross-lagged panel design: A review. *Perceptual and Motor Skills*, *45*, 883–893.

Marsh, A. A., & Blair, R. J. R. (2008). Deficits in facial affect recognition among antisocial populations: a meta-analysis. *Neuroscience & Biobehavioral Reviews*, *32*, 454–465. http://doi.org/10.1016/j.neubiorev.2007.08.003

Marsh, A. A., Finger, E. C., Mitchell, D. G. V, Reid, M. E., Sims, C., Kosson, D. S., … Blair, R. J. R. (2008). Reduced amygdala response to fearful expressions in children and adolescents with callous-unemotional traits and disruptive behavior disorders. *American Journal of Psychiatry*, *165*, 712–720. http://doi.org/appi.ajp.2007.07071145

[pii]\r10.1176/appi.ajp.2007.07071145

Masten, A. S., Coatsworth, J. D., Neemann, J., Gest, S. D., Tellegen, A., & Garmezy, N. (1995). The structure and coherence of competence from childhood through adolescence. *Child Development*, *66*, 1635–1659. http://doi.org/0009-3920/95/6606-0012

Maxwell, S. E., Cole, D. A., & Mitchell, M. A. (2011). Bias in cross-sectional analyses of longitudinal mediation: Partial and complete mediation under an autoregressive model. *Multivariate Behavioral Research*, *46*, 816–841. http://doi.org/10.1080/00273171.2011.606716

McQuade, J. D., Murray-Close, D., Shoulberg, E. K., & Hoza, B. (2013). Working memory and social functioning in children. *Journal of Experimental Child Psychology*, *115*, 422–435. http://doi.org/10.1016/j.jecp.2013.03.002

Merriam-Webster. (2018). *Merriam-Webster*. Retrieved from https://www.merriam-webster.com/dictionary/antisocial

Miller, S. A. (2009). Children's understanding of second-order mental states. *Psychological Bulletin*, *135*, 749–773. http://doi.org/10.1037/a0016854

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, *78*, 622–646. http://doi.org/10.1111/j.1467-8624.2007.01018.x

Miner, J. L., & Clarke-Stewart, K. A. (2008). Trajectories of externalizing behavior from age 2 to age 9: relations with gender, temperament, ethnicity, parenting, and rater. *Developmental Psychology*, *44*, 771–786. http://doi.org/10.1037/0012-1649.44.3.771

Mitchell, D. G. V., Colledge, E., Leonard, A., & Blair, R. J. R. (2002). Risky decisions and response reversal: Is there evidence of orbitofrontal cortex dysfunction in psychopathic individuals? *Neuropsychologia*, *40*, 2013–2022. http://doi.org/10.1016/S0028-3932(02)00056-8

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. http://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “Frontal Lobe” tasks: a latent variable analysis. *Cognitive Psychology*, *41*, 49–100. http://doi.org/10.1006/cogp.1999.0734

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. *Psychological Review*, *100*, 674–701. http://doi.org/10.1037/0033-295X.100.4.674

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., … Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 2693–2698. http://doi.org/10.1073/pnas.1010076108

Moffitt, T. E., & Caspi, A. (2001). Childhood predictors differentiate life-course persistent and adolescence-limited antisocial pathways among males and females. *Development and Psychopathology*, *13*, 355–375. http://doi.org/10.1017/S0954579401002097

Moffitt, T. E., Caspi, A., Dickson, N., Silva, P., & Stanton, W. (1996). Childhood-onset versus adolescent-onset antisocial conduct problems in males: Natural history from ages 3 to 18 years. *Development and Psychopathology*, *8*, 399–424. Retrieved from http://journals.cambridge.org

Moffitt, T. E., Caspi, A., Rutter, M., & Silva, P. A. (2001). *Sex differences in antisocial behaviour - Conduct disorder, delinquency, and violence in the Dunedin longitudinal study*. Cambridge, UK: Cambridge University Press.

Montroy, J. J., Bowles, R. P., & Skibbe, L. E. (2016). The effect of peers’ self-regulation on preschooler’s self-regulation and literacy growth. *Journal of Applied Developmental Psychology*, *46*, 73–83. http://doi.org/10.1016/j.appdev.2016.09.001

References

Morgan, A. B., & Lilienfeld, S. O. (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

Moses, L. J., & Tahiroglu, D. (2010). Clarifying the relation between executive function and children's theories of mind. In B. W. Sokol, U. Müller, J. I. M. Carpendale, A. R. Young, & G. Iarocci (Eds.), *Self and social regulation: Social interaction and the development of social understanding and executive functions* (pp. 218–233). New York, NY: Oxford University Press. http://doi.org/10.1093/acprof:oso/9780195327694.003.0009

Murray-Close, D., Ostrov, J. M., & Crick, N. R. (2007). A short-term longitudinal study of growth of relational aggression during middle childhood: Associations with gender, friendship intimacy, and internalizing problems. *Development and Psychopathology*, *19*, 187–203. http://doi.org/10.1017/S0954579407070101

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide. Statistical analysis with latent variables* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Nagin, D., & Tremblay, R. E. (1999). Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. *Child Development*, *70*, 1181–1196. http://doi.org/10.1111/1467-8624.00086

O'Toole, S. E., Monks, C. P., & Tsermentseli, S. (2017). Executive function and theory of mind as predictors of aggressive and prosocial behavior and peer acceptance in early childhood. *Social Development*, *26*, 907–920. http://doi.org/10.1111/sode.12231

Odgers, C. L., Moffitt, T. E., Broadbent, J. M., Dickson, N., Hancox, R. J., Harrington, H., … Caspi, A. (2008). Female and male antisocial trajectories: from childhood origins to adult outcomes. *Development and Psychopathology*, *20*, 673–716. http://doi.org/10.1017/S0954579408000333

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*, 1063–1107. http://doi.org/10.1111/j.1745-9125.2011.00252.x

Olson, S. L., Choe, D. E., & Sameroff, A. J. (2017). Trajectories of child externalizing problems between ages 3 and 10 years: Contributions of children's early effortful control, theory of mind, and parenting experiences. *Development and Psychopathology*, *29*, 1333–1351. http://doi.org/10.1017/S095457941700030X

Olson, S. L., Sameroff, A. J., Kerr, D. C. R., Lopez, N. L., & Wellman, H. M. (2005). Developmental foundations of externalizing problems in young children: The role of effortful control. *Development and Psychopathology*, *17*, 25–45. http://doi.org/10.1017/s0954579405050029

Olweus, D. (1999). *Bullying at school: What we know and what we can do.* Oxford: Blackwell Publishers.

Onishi, K., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*, 255–258. Retrieved from www.sciencemag.org

Ostrov, J. M., Murray-Close, D., Godleski, S. A., & Hart, E. J. (2013). Prospective associations between forms and functions of aggression and social and affective processes during early childhood. *Journal of Experimental Child Psychology*, *116*, 19–36. http://doi.org/10.1016/j.jecp.2012.12.009

Overman, W. H. (2004). Sex differences in early childhood, adolescence, and adulthood on cognitive tasks that rely on orbital prefrontal cortex. *Brain and Cognition*, *55*, 134–147. http://doi.org/10.1016/S0278-2626(03)00279-3

Pasterski, V., Golombok, S., & Hines, M. (2011). Sex differences in social behavior. In P. Smith & C. Hart (Eds.), *Blackwell handbook of childhood social development* (2nd ed., pp. 281–298). Oxford, England: Blackwell.

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *37*(I), 51–87. http://doi.org/10.1111/j.1469-7610.1996.tb01380.x

References

Perlman, S. B., Jones, B. M., Wakschlag, L. S., Axelson, D., Birmaher, B., & Phillips, M. L. (2015). Neural substrates of child irritability in typically developing and psychiatric populations. *Developmental Cognitive Neuroscience*, *14*, 71–80. http://doi.org/10.1016/j.dcn.2015.07.003

Perner, J., Lang, B., & Kloo, D. (2002). Theory of mind and self-control: more than a common problem of inhibition. *Child Development*, *73*, 752–67. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12038549

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that..." attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, *39*, 437–471. http://doi.org/10.1016/0022-0965(85)90051-7

Petermann, F., & Petermann, U. (Eds. . (2007). *Hamburg-Wechsler-Intelligenztest für Kinder - IV (HAWIK-IV) [Wechsler intelligence scale for children IV].* Göttingen, Deutschland: Hogrefe.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *The Journal of Applied Psychology*, *88*, 879–903. http://doi.org/10.1037/0021-9010.88.5.879

Poland, S. E., Monks, C. P., & Tsermentseli, S. (2016). Cool and hot executive function as predictors of aggression in early childhood: Differentiating between the function and form of aggression. *British Journal of Developmental Psychology*, *34*, 181–197. http://doi.org/10.1111/bjdp.12122

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*, 183–193. http://doi.org/10.1007/s10802-008-9266-0

Polman, H., de Castro, B., 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*, 522–535. http://doi.org/10.1007/s10802-007-9109-4

Preacher, K. J. (2015). Advances in mediation analysis: A survey and synthesis of new developments. *Annual Review of Psychology*, *66*, 825–852. http://doi.org/10.1146/annurev-psych-010814-015258

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behavioral and Brain Sciences*, *4*, 515–526. Retrieved from http://journals.sagepub.com

Preskorn, H., & Baker, B. (2002). The overlap of DSM-IV syndromes: Potential implications for the practice of polypsychopharmacology, psychiatric drug development and the human genome project. *Journal of Psychiatric Practice*, *8*, 170–177.

Putnam, S. P., Rothbart, M. K., & Gartstein, M. A. (2008). Homotypic and heterotypic continuity of fine-grained temperament during infancy, toddlerhood, and early childhood. *Infant and Child Development*, *17*, 387–405. http://doi.org/10.1002/icd.582

Raaijmakers, M. A. J., Smidts, D. P., Sergeant, J. A., Maassen, G. H., Posthumus, J. A., Van Engeland, H., & Matthys, W. (2008). Executive functions in preschool children with aggressive behavior: Impairments in inhibitory control. *Journal of Abnormal Child Psychology*, *36*, 1097–1107. http://doi.org/10.1007/s10802-008-9235-7

Rathert, J., Fite, P. J., & Gaertner, A. E. (2011). Associations between effortful control, psychological control and proactive and reactive aggression. *Child Psychiatry and Human Development*, *42*, 609–621. http://doi.org/10.1007/s10578-011-0236-3

Ravens-Sieberer, U., Wille, N., Erhart, M., Bettge, S., Wittchen, H.-U., Rothenberger, A., … Döpfner, M. (2008). Prevalence of mental health problems among children and adolescents in Germany: results of the BELLA study within the National Health Interview and Examination Survey. *European Child & Adolescent Psychiatry*, *17*(S1), 22–33. http://doi.org/10.1007/s00787-008-1003-2

Reader, M. J., Harris, E. L., Schuerholz, L. J., & Denkla, M. B. (1994). Attention deficit hyperactivity disorder and executive dysfunction. *Developmental Neuropsychology*, *10*, 493–512. http://doi.org/10.1080/87565649409540598

Renouf, A., Brendgen, M., Parent, S., ..., & Séguin, J. R. (2010). Relations between theory of mind and indirect and physical aggression in kindergarten: Evidence of the moderating role of prosocial behaviors. *Social Development*, *19*, 535–555. http://doi.org/10.1111/j.1467-9507.2009.00552.x

Renouf, A., Brendgen, M., Séguin, J. R., Vitaro, F., Boivin, M., Dionne, G., … Pérusse, D. (2010). Interactive links between theory of mind, peer victimization, and reactive and proactive aggression. *Journal of Abnormal Child Psychology*, *38*, 1109–1123. http://doi.org/10.1007/s10802-010-9432-z

Rhee, S. H., & Waldman, I. D. (2002). Genetic and environmental influences on antisocial behavior: A meta-analysis of twin and adoption studies. *Psychological Bulletin*, *128*, 490–529. http://doi.org/10.1037/0033-2909.128.3.490

Rhoades, B. L., Greenberg, M. T., & Domitrovich, C. E. (2009). The contribution of inhibitory control to preschoolers' social-emotional competence. *Journal of Applied Developmental Psychology*, *30*, 310–320. http://doi.org/10.1016/j.appdev.2008.12.012

Richell, R. A., Mitchell, D. G. V, Newman, C., Leonard, A., Baron-Cohen, S., & Blair, R. J. R. (2003). Theory of mind and psychopathy. *Neuropsychologia*, *41*(5), 523–526. http://doi.org/10.1016/S0028-3932

Romine, C. B., & Reynolds, C. R. (2005). A model of the development of frontal lobe functioning: Findings from a meta-analysis. *Applied Neuropsychology*, *12*(4), 190–201. http://doi.org/10.1207/s15324826an1204

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2011). Modeling a cascade of effects: The role of speed and executive functioning in preterm/full-term differences in academic achievement. *Developmental Science*, *14*, 1161–1175. http://doi.org/10.1111/j.1467-7687.2011.01068.x

Rothbart, M. K., & Rueda, M. R. (2005). The development of effortful control. In U. Mayr, E.

Awh, & S. Keele (Eds.), *Developing individuality in the human brain: A tribute to Michael I. Posner* (pp. 167–188). Washington, DC: American Psychological Association.

Röthlisberger, M., Neuenschwander, R., Michel, E., & Roebers, C. M. (2010). Exekutive Funktionen: Zugrundeliegende kognitive Prozesse und deren Korrelate bei Kindern im späten Vorschulalter [Executive functions: Underlying cognitive processes and their correlates in late preschool children]. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie*, *42*, 99–110. http://doi.org/10.1026/0049-8637/a000010

Ruffman, T. (2014). To belief or not belief: Children's theory of mind. *Developmental Review*, *34*, 265–293. http://doi.org/10.1016/j.dr.2014.04.001

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514. http://doi.org/10.1007/BF02296192

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research - Online*, *8*(2), 23–74. Retrieved from https://www.dgps.de/fachgruppen/methoden/mpr-online/

Schoemaker, K., Mulder, H., Deković, M., & Matthys, W. (2013). Executive functions in preschool children with externalizing behavior problems: A meta-analysis. *Journal of Abnormal Child Psychology*, *41*, 457–471. http://doi.org/10.1007/s10802-012-9684-x

Scott, S., Knapp, M., Henderson, J., & Maughan, B. (2001). Financial cost of social exclusion: follow-up study of antisocial children into adulthood. *British Medical Journal (Clinical Research Ed.)*, *323*(7306), 191–194. http://doi.org/10.1136/bmj.323.7306.191

Sebastian, C. L., Fontaine, N. M. G., Bird, G., Blakemore, S.-J., De Brito, S. A., McCrory, E. J. P., & Viding, E. (2012). Neural processing associated with cognitive and affective theory of mind in adolescents and adults. *Social Cognitive and Affective Neuroscience*, *7*(1), 53–63. http://doi.org/10.1093/scan/nsr023

Sebastian, C. L., McCrory, E. J. P., Cecil, C. a M., Lockwood, P. L., De Brito, S. a, Fontaine, N.

M. G., & Viding, E. (2012). Neural responses to affective and cognitive theory of mind in children with conduct problems and varying levels of callous-unemotional traits. *Archives of General Psychiatry*, *69*, 814–822. http://doi.org/10.1001/archgenpsychiatry.2011.2070

Séguin, J. R. (2009). The frontal lobe and aggression. *European Journal of Developmental Psychology*, *6*, 100–119. http://doi.org/10.1080/17405620701669871

Séguin, J. R., Arseneault, L., & Tremblay, R. E. (2007). The development of self-regulation: Toward the integration of cognition and emotion. *Cognitive Development*, *22*, 401–405. http://doi.org/10.1016/j.cogdev.2007.08.008

Séguin, J. R., & Zelazo, P. D. (2005). Executive Function in early physical aggression. In R. E. Tremblay, W. W. Hartup, & J. Archer (Eds.), *Developmental origins of aggression* (pp. 307–329). New York: Guilford Press.

Sellers, R., Maughan, B., Pickles, A., Thapar, A., & Collishaw, S. (2015). Trends in parent- and teacher-rated emotional, conduct and ADHD problems and their impact in prepubertal children in Great Britain: 1999-2008. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *56*(1), 49–57. http://doi.org/10.1111/jcpp.12273

Shamay-Tsoory, S. G., & Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: A lesion study. *Neuropsychologia*, *45*, 3054–3067. http://doi.org/10.1016/j.neuropsychologia.2007.05.021

Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain : A Journal of Neurology*, *132*(Pt 3), 617–627. http://doi.org/10.1093/brain/awn279

Shamay-Tsoory, S. G., Harari, H., Aharon-Peretz, J., & Levkovitz, Y. (2010). The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. *Cortex*, *46*, 668–677. http://doi.org/10.1016/j.cortex.2009.04.008

Shamay-Tsoory, S. G., Tibi-Elhanany, Y., & Aharon-Peretz, J. (2006). The ventromedical prefrontal cortex is involved in understanding affective but not cognitive theory of mind stories. *Social Neuroscience*, *1*(3–4), 149–166. http://doi.org/10.1080/17470910600985589

Shamay-Tsoory, S. G., Tomer, R., Berger, B. D., & Aharon-Peretz, J. (2003). Characterization of empathy deficits following prefrontal brain damage: The role of the right ventromedial prefrontal cortex. *Journal of Cognitive Neuroscience*, *15*, 324–337. http://doi.org/10.1162/089892903321593063

Sharp, C. (2008). Theory of mind and conduct problems in children: Deficits in reading the “emotions of the eyes.” *Cognition and Emotion*, *22*, 1149–1158. http://doi.org/10.1080/02699930701667586

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*, 422–445. http://doi.org/10.1037/1082-989X.7.4.422

Simonds, J. (2006). *The role of reward sensitivity and response: Execution in childhood extraversion*. Unpublished doctoral dissertation. University of Oregon.

Snyder, J. (2015). Coercive family processes in the development of externalizing behavior: Incorporating neurobiology into intervention research. In T. P. Beauchaine & S. P. Hinshaw (Eds.), *The Oxford Handbook of Externalizing Spectrum Disorders* (pp. 286–302). New York, NY: Oxford University Press. http://doi.org/10.1093/oxfordhb/9780199324675.013.009

Snyder, J., Schrepferman, L., McEachern, A., & Suarez, M. (2010). Early covert conduct problems: Phenomenology, prevalence, cross-cetting diffusion, growth and consequences. *The Journal of Behavior Analysis of Offender and Victim Treatment and Prevention*, *2*(1), 4–19. http://doi.org/10.1037/h0100471

Song, J.-H., Volling, B. L., Lane, J. D., & Wellman, H. M. (2016). Aggression, sibling antagonism, and theory of mind during the first year of siblinghood: A developmental cascade model. *Child Development*, *87*, 1250–1263. http://doi.org/10.1111/cdev.12530

References

Spieker, S. J., Campbell, S. B., Vandergrift, N., Pierce, K. M., Cauffman, E., Susman, E. J., & Roisman, G. I. (2012). Relational aggression in middle childhood: Predictors and adolescent outcomes. *Social Development*, *21*, 354–375. http://doi.org/10.1111/j.1467-9507.2011.00631.x

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology*, *59*, 745–759. http://doi.org/10.1080/17470210500162854

Steinberg, L., & Morris, A. S. (2001). Adolescent development. *Annual Review of Psychology*, *52*, 83–110. http://doi.org/0066-4308/01/0201

Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology*, *30*, 395–402. http://doi.org/10.1037/0012-1649.30.3.395

Sutton, J., Smith, P. K., & Swettenham, J. (1999). Bullying and "theory of mind": A critique of the "social skills deficit" view of anti-social behaviour. *Social Development*, *8*, 117–127. http://doi.org/10.11111467-9507.00083

Sutton, J., Smith, P. K., & Swettenham, J. (1999). Social cognition and bullying: Social inadequacy or skilled manipulation? *British Journal of Developmental Psychology*, *17*, 435–450. http://doi.org/10.1348/026151099165384

Teufel, C., Fletcher, P. C., & Davis, G. (2010). Seeing other minds: attributed mental states influence perception. *Trends in Cognitive Sciences*, *14*, 376–382. http://doi.org/10.1016/j.tics.2010.05.005

Tye, C., Bedford, R., Asherson, P., Ashwood, K. L., Azadi, B., Bolton, P., & McLoughlin, G. (2017). Callous-unemotional traits moderate executive function in children with ASD and ADHD: A pilot event-related potential study. *Developmental Cognitive Neuroscience*, *26*, 84–90. http://doi.org/10.1016/j.dcn.2017.06.002

Vaillancourt, T., Miller, J. L., Fagbemi, J., Côté, S. M., & Tremblay, R. E. (2007). Trajectories and predictors of indirect aggression: Results from a nationally representative longitudinal study of canadian children aged 2–10. *Aggressive Behaviour*, *33*, 314–326. http://doi.org/10.1002/ab.20202

Valeri, L., & VanderWeele, T. J. (2013). Mediation analysis allowing for exposure-mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros. *Psychological Methods*, *18*, 137–150. http://doi.org/10.1037/a0031034

van den Bos, R., Homberg, J., & de Visser, L. (2013). A critical review of sex differences in decision-making tasks: Focus on the Iowa Gambling Task. *Behavioural Brain Research*, *238*(1), 95–108. http://doi.org/10.1016/j.bbr.2012.10.002

Van Nieuwenhuijzen, M., Van Rest, M. M., Embregts, P. J. C. M., Vriens, A., Oostermeijer, S., Van Bokhoven, I., & Matthys, W. (2017). Executive functions and social information processing in adolescents with severe behavior problems. *Child Neuropsychology*, *23*, 228–241. http://doi.org/10.1080/09297049.2015.1108396

Vetter, N. C., Altgassen, M., Phillips, L., Mahy, C. E. V, & Kliegel, M. (2013). Development of affective theory of mind across adolescence: disentangling the role of executive functions. *Developmental Neuropsychology*, *38*, 114–125. http://doi.org/10.1080/87565641.2012.733786

Vetter, N. C., Leipold, K., Kliegel, M., Phillips, L. H., & Altgassen, M. (2013). Ongoing development of social cognition in adolescence. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *19*, 615–629. http://doi.org/10.1080/09297049.2012.718324

Vetter, N. C., Weigelt, S., Döhnel, K., Smolka, M. N., & Kliegel, M. (2014). Ongoing neural development of affective theory of mind in adolescence. *Social Cognitive and Affective Neuroscience*, *9*, 1022–1029. http://doi.org/10.1093/scan/nst081

Villanueva Badenes, L., Clemente Estevan, R. A., & Garcia Bacete, F. J. (2000). Theory of Mind and Peer Rejection at School. *Social Development*, *9*, 271–283. http://doi.org/10.1111/1467-9507.00125

Vitaro, F., Brendgen, M., & Barker, E. D. (2006). Subtypes of aggressive behaviors: A developmental perspective. *International Journal of Behavioral Development*, *30*, 12–19. http://doi.org/10.1177/0165025406059968

Vitaro, F., Brendgen, M., & Tremblay, R. E. (2002). Reactively and proactively aggressive children: Antecedent and subsequent characteristics. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *43*, 495–505. http://doi.org/10.1111/1469-7610.00040

Vloet, T. D., Großheinrich, N., Konrad, K., Freitag, C., & Herpertz-Dahlmann, B. (2014). Mädchen mit Störungen des Sozialverhaltens [Female conduct disorders]. *Zeitschrift Für Kinder- Und Jugendpsychiatrie Und Psychotherapie*, *42*(2), 95–108. http://doi.org/10.1024/1422-4917/a000276

Wakschlag, L. S., Estabrook, R., Petitclerc, A., Henry, D., Burns, J. L., Perlman, S. B., … Briggs-Gowan, M. L. (2015). Clinical Implications of a Dimensional Approach: The Normal/Abnormal Spectrum of Early Irritability. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*, 626–634. http://doi.org/https://doi.org/10.1016/j.jaac.2015.05.016

Walters, G. D. (2005). Proactive and reactive aggression: A lifestyle view. In J. P. Morgan (Ed.), *Psychology of aggression* (pp. 29–43). Hauppauge, NY: Nova Science Publishers.

Wang, Z., Devine, R. T., Wong, K. K., & Hughes, C. (2016). Theory of mind and executive function during middle childhood across cultures. *Journal of Experimental Child Psychology*, *149*, 6–22. http://doi.org/10.1016/j.jecp.2015.09.028

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*, 655–684.

Wellman, H. M., Phillips, A. T., Dunphy-Lelii, S., & LaLonde, N. (2004). Infant social attention predicts preschool social cognition. *Developmental Science*, *7*, 283–288.

http://doi.org/10.1111/j.1467-7687.2004.00347.x

Welsh, M. C., Pennington, B. F., & Groisser, D. B. (1991). A normative-developmental study of executive function: A window on prefrontal function in children. *Developmental Neuropsychology*, *7*, 131–149. http://doi.org/10.1080/87565649109540483

Welsh, M., & Peterson, E. (2014). Issues in the conceptualization and assessment of hot executive functions in childhood. *Journal of the International Neuropsychological Society*, *20*, 152–156. http://doi.org/10.1017/S1355617713001379

White, B. A., Jarrett, M. A., & Ollendick, T. H. (2013). Self-regulation deficits explain the link between reactive aggression and internalizing and externalizing behavior problems in children. *Journal of Psychopathology and Behavioral Assessment*, *35*, 1–9. http://doi.org/10.1007/s10862-012-9310-9

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103–128. http://doi.org/10.1016/0010-0277(83)90004-5

Zadeh, Y. Z., Im-Bolter, N., & Cohen, N. J. (2007). Social cognition and externalizing psychopathology: An investigation of the mediating role of language. *Journal of Abnormal Child Psychology*, *35*(2), 141–152. http://doi.org/10.1007/s10802-006-9052-9

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): a method of assessing executive function in children. *Nature Protocols*, *1*, 297–301. http://doi.org/10.1038/nprot.2006.46

Zelazo, P. D., & Anderson, J. E. (2013). What is cognitive control ? In P. D. Zelazo & M. D. Sera (Eds.), *Minnesota Symposia on Child Psychology: Developing Cognitive Control Processes: Mechanisms, Implications, and Interventions,* (Vol. 37, pp. 1–20). Hoboken, NJ: John Wiley & Sons. http://doi.org/10.1002/9781118732373

Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, *6*, 354–360.

http://doi.org/10.1111/j.1750-8606.2012.00246.x

Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology*, *1*(2), 198–226. http://doi.org/10.1037/1089-2680.1.2.198

Zelazo, P. D., & Müller, U. (2011). Executive function in typical and atypical development. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (2nd ed., pp. 574–603). Chichester, UK: Whiley-Blackwell. http://doi.org/10.1002/9781444325485.ch22

Zelazo, P. D., Qu, L. I., & Kesek, A. C. (2010). Hot executive function: Emotion and the development of cognitive control. In S. D. Calkins & M. A. Bell (Eds.), *Child development at the intersection of emotion and cognition* (pp. 97–111). Washington, DC: American Psychological Association. http://doi.org/10.1037/12059-006

Zimmermann, P., Gondan, M., & Fimm, B. (2002). KITAP Testbatterie zur Aufmerksamkeitsprüfung für Kinder [KITAP Attention Test Battery for Children]. Herzogenrath, Germany: Psychologische Testsysteme.

Erklärung

Hiermit versichere ich, die Dissertationsarbeit „Deficits in Theory of Mind and Executive Function as Risk Factors for Conduct Problems from Middle Childhood to Early Adolescence – a Longitudinal Perspective“ selbstständig angefertigt zu haben. Die Arbeit wird zur Promotion im Fach Psychologie eingereicht und ist selbstständig und ohne unzulässige Hilfe Dritter verfasst worden. Bei der Abfassung wurden nur die in der Dissertation angegebenen Hilfsmittel benutzt sowie alle wörtlich oder inhaltlich übernommenen Stellen als solche gekennzeichnet. Die Dissertation ist in der gegenwärtigen oder einer anderen Fassung in keinem früheren Promotionsverfahren angenommen oder abgelehnt worden. Ich habe an keiner anderen Hochschule ein Promotionsverfahren eröffnet.

I hereby declare that the dissertation "Deficits in Theory of Mind and Executive Function as Risk Factors for Conduct Problems from Middle Childhood to Early Adolescence – a Longitudinal Perspective" was composed by me. This document is being submitted for a doctorate in psychology and has been written independently and without undue assistance from third parties. In writing this document, only the resources indicated in the dissertation were used; and all representations of other work, verbatim or in substance, are marked as such. The dissertation has not been accepted or rejected in the current or in a different version in any other promotion procedures, nor have I applied for a promotion procedure at any other university.

Potsdam, April 2018 Anna Katharina Holl

Unterschrift