THEORY OF MIND DEVELOPMENT AND MORAL JUDGMENT AS DIFFERENTIAL PREDICTORS OF AGGRESSIVE AND PROSOCIAL BEHAVIORS IN A NORMATIVE PRESCHOOL SAMPLE

Erin Ruth Baker

A Dissertation
Submitted to the Graduate College of Bowling Green State University in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2016

Committee:
Marie S. Tisak, Advisor
Susannah Cleveland
Graduate Faculty Representative
John Tisak
Carolyn Tompsett
ABSTRACT

Marie S. Tisak, Advisor

The present study examined the relations between Theory of Mind and moral development as they impact specific aggressive behaviors and prosociality in preschoolers. Social-Cognitive and Social-Cognitive Domain Theory suggest that behavior should be considered a function of personal factors and environment, whereby each of these three facets impact and interact with one another, especially during the development of moral understanding. The current conceptualized hypothesized that moral judgment would predict socio-moral cognition, and – separately – that inhibitory control would predict Theory of Mind; additionally, a moderation of socio-moral cognition was posed on the expected relation from Theory of Mind to social outcomes. One hundred seventy-six preschoolers (Mₐ=53.3 months) completed measures of inhibitory control, moral understanding, morally imbedded Theory of Mind, verbal skills, and a traditional Theory of Mind battery consisting of five tasks. In addition, teacher- and child self-report data were collected for four specific aspects of aggressive behavior, and prosociality. The proposed model was a good fit to the data using teacher-report data. Results indicate that for children high in mental state understanding, applied moral cognitive development influences teacher-rated prosociality, such that those who were better able to understand another’s motive tended to be less prosocial—for those without motive understanding, there was no relation from ToM on prosocial behavior. Implications are discussed in regards to theoretical development and previous empirical findings.
For Molly, Jake, Stephanie, Wrenn, Chapel, Fischer, Claire, Christian, Henry, Harrison, and all future tiny humans who constantly provide endless inspiration.

And for the late Ruth Baker, and the late Stan Goodwin, who were always patient, understanding, and loving.
ACKNOWLEDGMENTS

First, I must extend my thanks to my doctoral advisor, Dr. Marie Tisak, for her continued assistance and guidance. I would also like to acknowledge Dr. John Tisak and Dr. Carolyn Tompsett for their comments and suggestions on both my dissertation and preliminary project—without their help, this study would not have been possible. I must also thank my outside committee members, Dr. Rachel Reinhart and Dr. Susannah Cleveland, for their helpful perspectives and contributions. I would also like to thank Dr. Shari Kidwell and Dr. Steve Hupp, for their contagious enthusiasm of child development.

Rena Onady, Samantha Awada, and Michael Hostetter must be recognized as well, as they were the most responsible, reliable, and mature research assistants one could ask for. Their dedication to this study was paramount to the success and quick finish of the data collection process.

Of course, I would like to thank my family, for supporting me even though they still have no idea what I do. To Jessie, I am officially no longer a “college student”. And I must acknowledge my friends, who provided endless support even when they knew exactly when I do but thought it was boring.

I would like to thank the schools administrators, teachers, parents, and children who participated in this study. Without their trust and confidence in me, and this study, I would not have been able to accomplish this dream.

Lastly, I would like to thank Dr. Dev Dalal. For all the things.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter I: Introduction</td>
<td>Social-Cognitive and Social-Cognitive Domain Theories</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Theory of Mind</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Moral Judgment and Theory of Mind Development</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Inhibitory Control</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Aggression</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Physical Aggression</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Relational Aggression</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Proactive Aggression</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Reactive Aggression</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Inhibitory Control and Aggression</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Prosocial Behavior</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Theory of Mind and Prosociality</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Current Study</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Hypotheses</td>
<td>24</td>
</tr>
<tr>
<td>Chapter II: Methods</td>
<td>Participants</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Procedures</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Measures</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Demographic information</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Theory of Mind Battery</td>
<td>29</td>
</tr>
</tbody>
</table>
Diverse Desires Task. ................................................................. 30
Knowledge Access Task ............................................................... 30
Appearance-Reality Task. ............................................................... 31
Explicit False-Belief Task. ............................................................. 31
Real-Apparent Emotion Task. ......................................................... 32
Moral Theory of Mind (MoToM) Task .................................................. 33
Moral Transgression Task ............................................................... 36
Inhibitory Control Task ................................................................. 37
Aggression ....................................................................................... 38
Aggressive Behavior Questionnaire – Teacher Report ............... 38
Aggressive Behaviors – Child Interview ........................................ 39
Prosocial Behaviors ........................................................................ 40
Child Behavior Scale – Teacher Report ........................................ 40
Prosocial Behaviors – Child Interview .......................................... 41
Verbal Skills ................................................................................... 41
Peabody Picture Vocabulary Test – Version 4 .................................. 41
CHAPTER III: RESULTS .................................................................... 43
Hypothesis 1: Scale Reliability and Construction ....................... 43
Theory of Mind .............................................................................. 43
Aggression ....................................................................................... 46
Teacher Report ............................................................................... 46
Child Report .................................................................................. 47
Prosocial Behaviors ....................................................................... 48
Teacher Report .................................................................................. 48
Child Report ...................................................................................... 48
Hypothesis 2: Effects of Age .............................................................. 48
Hypothesis 3: Effects of Gender ....................................................... 50
Hypothesis 4: Effects of Socio-economic status on Aggression .......... 50
Hypotheses 5-8: Relations among Primary Constructs of Interest ........ 51
   Path Analysis: Teacher Report ...................................................... 53
   Primary Sample ........................................................................ 53
   Sample Composition 1: MoToM Initial Restriction ...................... 53
   Sample Composition 2: MoToM Conservative Restriction .......... 54
   Sample Composition 3: Age Restriction .................................... 54
   Path Analysis: Child Self-Report .................................................. 54
   Primary Sample ........................................................................ 55
   Sample Composition 1: MoToM Initial Restriction ...................... 55
   Sample Composition 2: MoToM Conservative Restriction .......... 56
   Sample Composition 3: Age Restriction .................................... 56
Hypothesis 9: ToM, MoToM, and Behavior: Moderation Analyses ........ 58
CHAPTER IV: DISCUSSION ............................................................... 60
   Review of Findings and Contributions ........................................ 61
   Review of Individual Relations .................................................. 61
   Review of Proposed Model and Interpretation ............................ 63
   Applied Moral Cognition within Early Childhood ....................... 64
   Findings for Prosociality ............................................................. 67
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>96</td>
</tr>
<tr>
<td>2</td>
<td>97</td>
</tr>
<tr>
<td>3</td>
<td>98</td>
</tr>
<tr>
<td>4</td>
<td>99</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>6</td>
<td>101</td>
</tr>
<tr>
<td>7</td>
<td>102</td>
</tr>
<tr>
<td>8</td>
<td>103</td>
</tr>
<tr>
<td>9</td>
<td>104</td>
</tr>
</tbody>
</table>

1. Descriptive statistics and demographics by age.
2. Individual school sample size and demographics.
3. Theory of Mind battery.
4. Constructs and measures.
5. Demographic statistics for the five sample compositions.
6. Item measure summary and fit statistics for 5-item Rasch Model.
8. Effects of age on IC-accuracy, IC-reaction time, ToM, MoToM, and behavioral outcomes.
9. Means for variables of interest by gender within the conservative MoToM restricted sample.
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hypothesized relations between primary constructs</td>
<td>105</td>
</tr>
<tr>
<td>2</td>
<td>Graded scale used in MoToM and Moral Transgression Tasks.</td>
<td>106</td>
</tr>
<tr>
<td>3a</td>
<td>Night Stroop Task Card Image</td>
<td>107</td>
</tr>
<tr>
<td>3b</td>
<td>Day Stroop Task Card Images</td>
<td>107</td>
</tr>
<tr>
<td>4</td>
<td>Path model: Teacher Data</td>
<td>108</td>
</tr>
<tr>
<td>5</td>
<td>Path model: Child Data</td>
<td>109</td>
</tr>
<tr>
<td>6</td>
<td>Moderation of MoToM on the relation from ToM on teacher-reported prosocial behavior</td>
<td>110</td>
</tr>
</tbody>
</table>
THEORY OF MIND DEVELOPMENT AND MORAL JUDGMENT AS DIFFERENTIAL PREDICTORS OF AGGRESSIVE AND PROSOCIAL BEHAVIORS IN A NORMATIVE PRESCHOOL SAMPLE

Researchers in development have long speculated on the relationship between advances in metacognitive development and moral understanding, especially within the realm of social interactions. Specifically, much research bridging these areas focuses on how individuals judge the behaviors of others - whether or not the intentionality of actions influences the moral valence of those actions, for example. However, there exists a dearth of empirical evidence regarding how these developing internal processes influence one’s own behavior, specifically among preschool-aged children, an age range during which cognition, moral understanding, and social understanding make marked advances in normal development. The following discussion begins with a brief integration of work by Bandura (1986; 1991; 1999) and Turiel (1983; 2006; 2008) on Social-Cognitive and Social-Cognitive Domain Theories, respectively, followed by a summary of the existing literature on Theory of Mind (ToM), and the theoretical and empirically supported interrelations with morality and Moral Theory of Mind (MoToM). Executive functioning, namely inhibitory control, is then expounded upon in regard to ToM and moral development, which leads to a discussion on influences on aggressive behaviors. Finally, theoretical metacognitive and moral explanations for different aspects of aggressive behavior and relations with prosocial behavior are explained.

Social-Cognitive and Social-Cognitive Domain Theories

In explaining social behaviors in child development, several theories have arisen which, combined, offer invaluable insight into the development of social competence (Tisak, Tisak, & Goldstein, 2006). Bandura’s (1986; 1989; 1991) social-cognitive theory, while originally
centered in behaviorism, has been expanded to include several internal components. Specifically, later theoretical extensions are premised on a “triadic reciprocal causation,” which incorporates behaviors, thoughts, affective states, and other internal and environmental factors in predicting social behaviors, namely aggression (Bandura, 1989). By this understanding, children are not simply reactors to environmental stimuli, but are both mindful and purposeful in their reactions to such situations. Children accomplish this by using goal-focused understanding, becoming efficacious in regard to this goal, and regulating inappropriate responses via affective knowledge (Bandura, 1989; 1991). For example, if a child selects a particular goal (e.g, acquiring a toy that is being used by another boy), the child might think of possible scenarios by which he could accomplish that goal. The child might think that he could simply take the toy from the boy, but recognizes that this is not acceptable behavior. Instead, the child could more appropriately ask the target child if they could take turns or share the toy. In recognizing that he would be more successful in the latter scenario, the child develops his self-efficacy based upon previous successful interactions. Also, by entertaining multiple avenues by which to achieve his goal and selecting the most socially appropriate, the child is self-regulating his immediate impulses. Lastly, by extrapolating upon previous experiences and actions to future consequences, the child is again enhancing both his self-efficacy and self-regulatory skills (Bandura, 1989).

Turiel’s (1983; Smetana, 2006) theoretical explanation of social interactions is based on the premise that humans have several foci of social understanding, which he termed domains. Domain theory supposes that social behaviors and understanding are innately categorized into different groups: the moral domain, the conventional domain, and the personal domain (Tisak et al., 2006). The moral domain, of special focus here, regards the welfare of others and includes concepts such as harm and safety. The conventional and personal domains, less germane to this
study, refer to social rules of conduct designed at group coordination, and individual concerns such as choice and preference, respectively (Smetana, 2006; Tisak et al. 2006). The moral domain, according to Turiel (1983), can be explored and understood by children according to three possible scenarios: the child is the actor of the social action, the child is the recipient of the social action, or the child is simply a witness to the social action (Tisak et al.). For example, children can learn about moral judgments (the rightness/wrongness of an act) and behaviors (for example, aggression) by: behaving aggressively, being a victim of aggression, or witnessing aggression.

Moral knowledge, then, stems from an understanding of welfare and harm; however, this differs from moral reasoning, in which a person’s thoughts are embedded with their own emotions, and of which the person is an active participant (Turiel, 2006). Moral understanding, regarded as “obligatory, universally applicable, impersonal, and normatively binding” (p. 121, Smetana, 2006), refers to a sense of humanity and equality, rather than cultural or societal expectations (Helwig, Tisak, & Turiel, 1990; Turiel, 2006). Moral understanding, then, generally develops quite quickly, between 3- and 4-years of age, once children gain experiences with these concepts. Moral reasoning, on the other hand, is defined as, “thinking in dilemma situations where issues of justice, fairness, or caring are prevalent” (p. 565, Carlo, 2006), and concerns applying the moral knowledge in social interactions involving the self and others. Moral reasoning, then, is more nuanced to the individual’s beliefs and emotions, and becomes more sophisticated with age.

The two theories converge on the application of these respective understandings in that both theories stipulate an internal cognitive component, an affective evaluation, and lastly, a behavioral response. While Bandura theorizes more on personal emotional states, Turiel
acknowledges the inherent socially evaluative component necessary in morally-relevant interactions; in fact, multiple studies show that children as young as 3 acknowledge that immoral acts, such as hitting, are wrong even when permitted by authority figures (Nucci & Turiel, 1993; Smetana, Jambon, Conry-Murray, & Sturge-Apple, 2012; Tisak, 1986) and by rules (Smetana et al., 2006). This stands in contrast, for example, with young children’s understanding of conventional rules, which are viewed as arbitrary and flexible, and transgressions of which are seen as less punishable (Tisak et al., 2006).

**Theory of Mind**

Theory of Mind (ToM), a phenomenon that lies within the realm of social cognition, is generally the ability to understand another’s mental state, and to use this information in order to predict another person’s behaviors and expectations (Frye & Moore, 1991). ToM requires that the individual understands that another person is an independently thinking agent, or the individual has a theory that another person has a mind all his own. In some literature, (e.g., Hynes, Baird, & Grafton, 2006), ToM is conceptualized interchangeably with perspective-taking; although the two are certainly related, ToM is a more complex and dynamic understanding of another’s mind. In much of the cognitive development literature perspective-taking refers to the knowledge that another person has a visually different view than does oneself, and this is considered a very basic level of ToM (Howlin, Baron-Cohen, & Hadwin, 1999).

For informational purposes, early conceptualization by Baron-Cohen and colleagues (Baron-Cohen, 2001; Baron-Cohen, Leslie, & Frith, 1985; Howling, Baron-Cohen, & Hadwin, 1999) identifies that there are five levels of ToM; the first two levels refer to visual perspective-taking (Simple, and Complex, respectively; Barnes-Holmes, McHugh, & Barnes-Holmes, 2004) whereas the third, fourth, and fifth levels of ToM refer to acquiring knowledge from assumed,
non-visual, information; predicting another’s actions given a true – and therefore shared – knowledge base; and action prediction from a false knowledge base (or false belief), respectively (Barnes-Holmes, McHugh, & Barnes-Holmes, 2004). Therefore ToM is a more elaborate and elegant ability to know another’s full range of mental abilities – beliefs, emotions, intentions, etc. – whereas perspective-taking refers to information from the physical world (Baron-Cohen, 2001). Thus perspective-taking alone is not ToM. That is, one can be able to take another’s perspective without being able to interpret their behaviors or anticipate their emotions.

Of note, because the current study is primarily interested in the mental and moral components that predict specific social behaviors, and that previous studies show that 3- to 4-year-old children clearly understand the mental-physical distinction (Wellman & Estes, 1986), perspective-taking and the assessment of such will not be examined here. For the present purposes, ToM refers to one’s ability to understand another’s mental states, specifically beliefs, desires, and intentions. Encapsulated within ToM are not only the successful encoding of the linguistically relevant information – such as tone and inflection – but also perceiving contextual factors as well as idiosyncratic biases among other actors, such as preferences and choice (Baron-Cohen, 1988). This information allows individuals to excogitate and predict others’ behaviors and motives, enabling personal influence, which has shown to be valuable across culture and ethnicity (Liu, Wellman, Tardif, & Sabbagh, 2004).

The majority of ToM research aims to identify at what age a matured understanding of human behavior can be expressed (Laranjo, Bernier, Meins, & Carlson, 2010; Mao & Ning, 2004; Mounoud, 1996; Nguyen & Frye, 1999), as well as other factors, such as assisting ToM development (Flobbe, Verbrugge, Hendricks, & Krämer, 2008; Randell & Peterson, 2009), and influences on subsequent social interactions, including social competence, desire, intentionality,
and bullying (Gopnik, 1990; Mao & Ning, 2004; Nguyen & Frye, 1999; Shakoor et al., 2012). In normal development, ToM is believed to begin to develop largely between 3 and 5 years (Mao & Ning, 2004; Perner, 1991). Though partial ToM can be evinced in younger children, such as in the use of deception and pretend play (Kuhn, 1999), full ToM development is believed to take place in the early school years, at approximately 5 to 6 years of age in normal development (Moore & Furrow, 1991; Siegal & Varley, 2002), corresponding to brain maturation in the frontal lobe. Younger children, while capable of differentiating reality from imagination, are unable to discern that another’s beliefs do not mirror their own; that is, children seem to imbue beliefs and assertions with a certain impermeability, while understanding the ephemerality of pretend play (Kuhn).

Generally, evidence points to a gradual development of metacognitive processes throughout the preschool and early childhood years. While much of ToM has focused on a child’s ability to understand another’s beliefs, with classic examples of false-belief tasks (Baron-Cohen, Leslie, & Frith, 1985; Wellman & Liu, 2004; Wimmer & Perner, 1983) ever-present in the literature (Killen, Mulvey, Richardson, Jampol, & Woodward, 2011; Loureiro & Souza, 2013; Smetana et al., 2012), understanding ToM from purely informational or cognitive perspectives fails to give a full picture of this construct. Affective ToM, or the ability for one to understand another’s emotional states, merits consideration as well, especially in studies of social behavior (Carlo, 2006; Dvash & Shamay-Tsoory, 2014).

Although both cognitive and affective ToM require one to represent another’s mental states, cognitive ToM requires only knowledge- or belief-based understanding, whereas affective ToM requires emotional understanding, which can involve some nuance given the situation. Recent work examining both cognitive and affective components of ToM has provided the field
with exciting research, particularly in the domain of neuroimaging. For instance, a recent meta-analysis found that both cognitive and affective ToM tasks activate the medial prefrontal cortex and bilateral temporo-parietal regions (Schurz, Radua, Aichhorn, Richland, & Perner, 2014), which seem to represent a “core neural network” for mental state representation. Cognitive ToM tasks, however, uniquely activate the cuneus and precuneus, responsible for mental imagery, and the affective ToM tasks activated the inferior frontal gyrus, which is responsible for emotional identification and action planning, and other prefrontal cortical structures (Schlaffke, Lissek, Lenz, Juckel, Schultz, Tengenthoff, et al., 2015; Schurz, et al., 2014).

In further examination of cognitive and affective ToM, studies on empathy invariably describe the intuitive and theoretical bases for shared variance (Davis & Stone, 2003). It is true that empathy and ToM share commonalities – both are mental abilities that require and are defined by the characteristic of understanding the mental state of another. If one were to parse ToM to specify the emotional component – affective ToM – this conversation becomes even more confusing. However, Dvash and Shamay-Tsoory (2014) clarified the distinction nicely: affective ToM refers to the identification, knowledge of, and understanding the existence of another’s emotions, whereas empathy refers to the simulation and stimulation of emotional states within oneself stemming from the identification, knowledge of, and understanding the existence of another’s emotions. Empathy is an emotional response within an individual in response to a target’s emotional state, whereas affective ToM is the understanding and recognition of a target’s emotional state (Dvash & Shamay-Tsoory, 2014).

Extending the previous discussion on neural networks responsible for ToM, neuroimaging research on empathy and ToM has identified that these are, indeed, separate yet similar constructs. While ToM and empathy both employ use of medial prefrontal cortex (Dvash
& Shamay-Tsoory, 2014; Schurz, et al., 2014) empathy activates internal structures responsible for emotional experiences, namely the amygdala and insular (Dvash & Shamay-Tsoory, 2014). Combined with theoretical distinction and neuroimaging studies, it is clear that cognitive ToM, affective ToM, and empathy are related, yet separate, mental tools.

**Moral Judgment and Theory of Mind Development**

Affective ToM is closely related to moral understanding and moral reasoning and has been argued by theorists to be a necessary component in explaining relations between more traditional ToM studies and studies on social behaviors (Carlo, 2006), such as prosociality (actions performed with the intent of benefitting others) and various types of aggression (acts that cause harm to others). Specifically, affective ToM centers on understanding others’ emotional states, which facilitates other-centered behavior in more realistic and predictable ways than simply using cognitive ToM, as many social interactions require an emotional understanding – knowing that another person might be sad if his toy is missing, for example. Many argue that studying ToM without including socially relevant tasks is in fact misrepresenting the construct (Killen et al., 2011) – many ToM tasks access knowledge of behavior based on factual information (“where is she going to look for her ball”; Baron-Cohen et al., 1985), but fail to assess knowledge of emotional responses (“how will she feel when she finds out her ball is gone”; Killen et al., 2011).

In fact, many studies argue that moral judgment is reliant upon an existing basic cognitive ToM (Killen et al., 2011; Lane, Wellman, Olson, LaBounty, & Kerr, 2010; Loureiro & Souza, 2013; Wainryb & Brehl, 2006); one cannot be expected to understand another’s emotional state – arguably a more complex and dynamic human experience – if one cannot understand another’s non-affective beliefs regarding factual information. This theory mirrors
early supposition by Kohlberg and Turiel (1971) and others that a cognitive understanding of mental states is a prerequisite to understanding affective mental states. Indeed, recent research (Killen et al., 2011) points to the long-standing assumption that moral behavior is in fact morally relevant ToM (MoToM), and that young children might have difficulty in successfully applying ToM in morally-relevant social situations, as these interactions can be complex and unpredictable at times. Empirical work by Boerger (2015) supports this: in a sample of middle-childhood-aged participants, moral identity – the moral valence to which one prescribes – predicted moral behavior, but only for children high in ToM. Children who were unsuccessful in their ability to understand another’s mental state did not behave in any way related to their moral identity (Boerger, 2015).

Studies involving younger samples have found similar results. For example, Loureiro and Souza (2013) found that preschool-aged children were more competent at assessing motivational valence when actions were clearly intentional, rather than accidental, which the authors argue suggests a cognitive foundation. That is, when children could easily distinguish the intent – positive or negative – of the actor, they were better at determining whether or not it was positive or negative; when the actor did something by accident, children had a difficult time understanding this. In examining this further, Killen and her colleagues (Kristen, Sodian, Licata, & Killen, 2015) explicitly examined standard false-belief development and morally embedded false-belief. As expected, knowledge of moral intention developed later than understanding an actor’s mistaken cognition (false belief), implying that ToM does indeed develop earlier than morally-relevant ToM (MoToM). Neuroimaging studies of ToM and morality show that basic moral judgments based on outcome and intention (rather than those based solely on outcome),
depend on cognitive processes mediated by specific brain areas known to be involved in ToM development (Young, Cushman, Hauser, & Saxe, 2007).

Conversely, other studies have contended that moral judgment develops before metacognitive advances, regarding moral understanding as a vital component is internalizing another’s mental state (Knobe, 2003a; 2004; 2005). Knobe (2003a) has illustrated that, for instance, the valence of an outcome influences perceived intentionality. The changing of outcome in a scenario (e.g., A chairman stating “I don’t care if the program [hurts/helps] the environment, I just want to make as much profit as I can”) affected whether or not individuals believed the positive/negative outcome to have been done purposefully. That is, individuals judge the negative outcome (“…hurts the environment…”) to have been committed on purpose; however, if the outcome is positive, then individuals judge the outcome to be accidental. These and other findings (Knobe 2003b; 2004) suggest the presence of a “moral faculty” (Dwyer, 1999) that would use morally relevant information to produce outcome judgments of the intentionality of the act (Knobe, 2006).

Furthermore, there are still studies of both ToM and several aspects of moral judgment that indicate that these two socio-cognitive constructs seem to emerge simultaneously (Leslie, Knobe, & Cohen, 2006), and furthermore have bidirectional influences (Smetana et al., 2012). Smetana et al., for example, found that preschool children, over a 1-year-time period, made specific advances in both ToM and moral judgment predicated by earlier moral judgment and ToM abilities, respectively. Specifically, ToM abilities at Wave 1 and Wave 2 predicted a child’s later understanding of moral actions as independent of rules, of rule non-alterability, and of deserved punishment for immoral committance. Mirrored results were found for predicting ToM development from previous moral understanding; understanding the non-permissibility of moral
actions and that immoral acts remain as such independent of authority approval predicted ToM scores at Waves 2 and 3. In relation to explicit behavioral interactions, Suway, Degnan, Sussman, and Fox (2012) found that ToM understanding at age 3 was explained by negative peer interactions (namely, the use of grabbing, demanding, and rejection) at age 2; Suway et al. (2012) identified that children who exhibited greater negative peer behaviors at age 2 performed more poorly on a ToM assessment at age 3, which – given that ToM develops per normative brain maturation, as previously mentioned – suggests that behavioral interactions might offer the child an opportunity to practice and monitor their early developing ToM and MoToM abilities (Fink, Begeer, Hunt, & de Rosnay, 2014).

Fink, Begeer, Peterson, Slaughter, and de Rosnay (2015) found further evidence for this theory, in that poor ToM performance at age 5 predicted friendlessness at age 7. In fact children in this study (Fink, et al., 2015) who were deemed “chronically friendless” (reported no friends at both age 5 and age 7) had poorer ToM abilities than did all other children, including children who were rejected by most but reported having some mutual friendships. Therefore, research in this area seems to converge on the assumption that social interactions, of both a cognitive and affective nature, are requisite to advancing personal socio-cognitive development in that they provide the opportunity to practice these abilities (Carpendale & Lewis, 2004; Fink, et al., 2014). While some moral transgressions do not require mental state understanding, Perner (1991) suggests that these internal experiences allow the child to make assumptions about the external, representational world (“Being hit hurts me, so hitting hurts others”), therein helping to further develop a representational ToM.

However, the relationship and appropriate use of moral understanding and ToM may not be as clear as this in regard to actual peer relationships, rather than peer interactions, as empirical
research on friendship shows. For example, young children are more willing to justify moral transgressions if the transgressor is a friend (Dunn, Cutting, & Demetriou, 2000; Slomkowski & Killen, 1992). Older children exhibit this phenomenon as well, citing jest based on interpersonal relationships, rather than aggression, as the explanation (Tisak & Tisak, 1996). Alternatively, research by Boerger (2015) indicates a moderating relationship in explaining aggressive behavior. In this study, although ToM failed to directly explain aggressive behavior, there was a ToM-by-moral-identity interaction whereby moral identity predicted aggression, but only for those children who were higher in ToM abilities, indicating that moral understanding may facilitate aggressive and prosocial behavior but only for children adept at understanding other’s mental states. That is, it may be true that children are capable of selecting appropriate social behaviors in line with a given situation’s moral valence only when they have the ability to understand their peers’ mental states.

Looking forward, research should attempt to incorporate morally relevant ToM tasks (MoToM; Killen et al., 2011) in concert with traditional ToM assessments (i.e., false-belief-location: Sally-Anne; Baron-Cohen et al., 1985) and measures of moral judgment (Baird & Astington, 2004; Loureiro & Souza, 2013) in younger samples. Doing so would not only help to fill a still existing gap in the literature (Killen et al.), but would also be more appropriate given the interconnectedness of developing morality and emotional salience in social situations requiring metacognitive abilities (Boerger, 2015).

**Inhibitory Control**

Self-regulation, arguably, is one of the most important cognitive instruments when engaging in social interactions. A vast body of research suggests that ToM skill is related to the development of executive function (Carlson, Moses, & Breton, 2002; Hughes & Ensor, 2007;
2008) assisted largely by brain maturation taking place in the frontal lobe between the ages of 3 and 5 (Sabbagh, 2006). Executive function refers to “higher order cognitive [processes] that underpin flexible goal-directed behavior” (Hughes & Ensor, 2007, p. 1447). Executive functioning is therefore an amalgam of cognitive aspects including attention, working memory, and inhibitory control, and how these interact to influence how one encodes, retrieves, and recognizes environmental information. Specifically, inhibitory control (IC) is the concept of being able to restrain impulses that impede the goals of the task at hand. IC is necessary for ToM development in that the ability to ignore a prepotent response, such as one’s own beliefs, to express a less automatic but more accurate statement is a large component of metacognition (Hughes & Ensor, 2006). Further, executive functions associated with certain brain areas have been found to predict certain aspects of ToM development, which may influence the emergence and idiosyncratic expression of ToM (Ding & Li, 2004). Inhibitory control and verbal fluency have consistently been found to predict ToM (Ahmed, Godsey, Kesian, Patel, & Miller, 2007; Flobbe, et al., 2008; Hughes & Ensor, 2007), giving cause for verbal assessment to be included in studies of ToM and IC.

Research involving children’s ToM has been related to inhibitory control (IC) across a battery of tests (Carlson & Moses, 2001). For example, when using 10 different IC tasks and four different ToM tasks, older children performed significantly better on six of the IC tests, and in all of the ToM tasks compared with a younger cohort. Further, every ToM task was significantly related to every IC measure, lending further support to the hypothesized relationship between ToM development and IC skills. Looking further into these findings, researchers found that ToM scores were also significantly related to both mental conflict scores, in terms of accuracy, and delay scores, in response time, within the IC paradigms; children with higher ToM scores,
compared with children scoring low on ToM, responded faster and with greater accuracy during the IC battery. Conflict scores, however, better predicted ToM scores than did delay scores (Carlson & Moses). These findings are consistent with an earlier version of this study: children who answered correctly on an IC task, but not necessarily faster, were more likely to be successful on two different ToM tasks (Baker, Tisak, Graupensperger, & Jensen, 2015).

Lastly, although much research suggests that social understanding is possible at a young age (see previous sections), some debate exists regarding the ability to use this knowledge at a young age. That is, is it possible that very young children are mentally capable of understanding another’s mental states and their moral valence in social situations, but aren’t fully capable of expressing this understanding due to underdeveloped inhibitory control, wherein one’s prepotent (self-directed) understanding is expressed more readily? Again, the argument posited by Mounoud (1996) regarding at what age children are able to convey ToM explicates these issues. More recent research has shown that development of metacognitive expression is actually in line with the development of metacognitive ability: children can express ToM abilities as they acquire them, rather than having metacognitive skills but lacking the IC to express it (Carlson, Claxton, & Moses, 2015).

**Aggression**

Aggression, or hostile or violent actions or attitudes towards another, is not limited to humans but is more advanced in humans, especially as we mature (Loeber & Hay, 1997). In childhood, aggression first presents as biting, kicking, and pushing (Tremblay et al., 1999), but as our brains develop and our social interactions become more elegant, aggressive behaviors develop as well, becoming more concealed in motive, intention, and exhibition (Björkqvist, Lagerspetz, & Kaukiainen, 1992). In line with initial categorization by Buss (1961), and
empirical exploration by Crick and colleagues (e.g., 1997; Crick & Dodge, 1996; Crick et al., 2006), a discussion on the differences in the behavioral expression (i.e., physical versus relational), and cognitive causes (i.e., proactive versus reactive) of aggression follows.

**Physical Aggression**

Physical, or sometimes called direct, aggression does not, as the name might imply, only include aggressive physical contact. Verbal expressions, such as threats of physical harm and non-contact physical motions are also included in this conceptualization. Research on physical aggression suggests an age-by-gender interaction (Björkqvist, 1994; Crick & Grotpeter, 1995; Hay et al., 2011; Walton, Harris, & Davidson, 2009). For example, in a sample of 2- to 3-year-olds, girls were just as likely as boys to show physical aggression when another person instigated the conflict, but girls were also more likely to desist from aggressive behavior over time (Hay et al., 2011).

Gender differences have also been found when examining physical aggression in older children. In evaluating personal experiences with conflict in 4th- to 6th-graders, Walton et al. (2009) found that boys demonstrated more severe physical aggression than girls, but only when the victim was also a boy. Tisak, Nucci, and Jankowski (1996) also found that when males are the instigators of physical moral transgressions, they primarily target other males; however, female instigators tended to equally victimize both genders. Furthermore, gender differences have also been found regarding intelligence and aggression; low IQ scores have been related to a high incidence of direct aggression in girls, while no relationship existed for boys (Valles & Knutson, 2008). This may follow gender stereotypes of aggression, in that physical aggression is often more allowable for males than females, regardless of intelligence.
Social repercussions of physical aggression have been fairly well studied. Physical aggression is likely to result in ostracism by peers (Minde, 1992), and both mothers and peers respond more harshly to events of direct aggression than indirect, or relational, aggression, to be discussed next. Additionally, higher indices of direct aggression, but not indirect aggression, related to lower popularity. Research by Tisak and colleagues (Goldstein, Tisak, & Boxer, 2002; Tisak, 1993) on preschoolers’ personal views of both relational and physical aggression found that children believe that overtly aggressive responses happen more often than they should under provocation (Goldstein et al., 2002), supporting findings of physical aggression being less accepted even at a young age (Tisak, 1993; Tisak, & Turiel, 1988).

**Relational Aggression**

Relational aggression, a term coined by Crick, Grotpeter, and Bigbee (2002), but also sometimes called social aggression (e.g., Galen & Underwood, 1997; Paquette & Underwood, 1999), or indirect aggression (e.g., Björkqvist, 1994; Lagerspetz, Björkqvist, & Peltonen, 1988), refers to emotional or psychological hostility aimed at damaging peer relationships or another’s social standing. Relational aggression therefore necessitates the involvement of a social intermediary. The end result of this type of aggression is often more delayed and subtle in comparison with physical aggression (Gendreau & Archer, 2005). While relationally aggressive skills develop through the childhood years, this type of aggression does not become the most common display of aggression until after childhood (Björkqvist et al., 1992).

Gender differences in relational aggression vary across studies. Some research indicates that females tend to show more relationally aggressive behaviors than males (Bonica, Arnold, Fisher, & Zeljo, 2003; Rys & Bear, 1997; Valles & Knutson, 2008), while other studies have not found such a relationship (Tomada & Schneider, 1997; Underwood, 2003). A study of late
adolescents showed that, on average, males were more likely to commit both types of aggression, whereas females tended to only exhibit relational aggression (Salmivalli & Kaukiainen, 2004). Such contradictions may be evidence of a more complicated connection; in a study of young adults’ friendships and romantic relationships, for example, females were more likely than males to exhibit relational aggression, but only in the context of romantic relationships (Goldstein, 2011). Some research has identified that the victim’s gender may play a role in this as well. Tisak, Tisak, and Laurene (2012) found in a study of 7- to 12-year-olds that children tend to be nicer to same-gendered classmates than opposite-gender peers.

While indirect aggression is significantly related to psychological impairment, loneliness, and depression for males and females (Crick, 1997; Crick & Grotpeter, 1995; Johnson & Foster, 2005; Rys & Bear, 1997), relational aggression still seems to be the preferred form of exhibiting frustrations, as physical aggression is often not permitted in certain contexts. In using two different provocation situations (i.e., relationally aggressive and physically aggressive), Goldstein et al. (2002) found that preschoolers felt that a relationally aggressive response was more acceptable than either a verbally or physically aggressive response. Findings on social implications are mixed (Prinstein & Cillessen, 2003). For example, high levels of relational aggression have been linked to increased peer status (Nelson, Robinson, & Hart, 2005) and an increase in mutually reported friendships developed over time (Burr, Ostrov, Jansen, Cullerton-Sen, & Crick, 2005), but also to increased peer rejection (Crick, 1997). This suggests that those exhibiting relational aggression may be able to mend damaged bonds given enough time. Hegemonic popularity or perceived popularity, rather than actual likeability, may also be to blame for disparate findings (Vaillancourt, 2005).
Considering the opinions young children hold regarding appropriateness of different types of aggression (Goldstein et al., 2002), it is useful to understand the developmental trajectory of relational aggression and the psychological perspectives of this phenomenon taken by children. For example, children in middle-to-late childhood and early adolescence view parental and peer influences on relational aggression differently; parental opinion seems to have authority in regard to accepting or rejecting physical aggression, but not relational aggression, whereas peer influences are important factors in both situations (Goldstein & Tisak, 2006; Goldstein, Tisak, Persson, & Boxer, 2006).

Aside from the ways in which one chooses to express aggression, research has focused on the perceived source of the aggression. Is the aggressor instigating the aggressive interaction (i.e., proactive aggression), or is the aggression a consequence of another’s ill-mannered actions (i.e., reactive aggression)? A detailed discussion of these two types follows.

**Proactive Aggression**

Proactive, also called instrumental, aggression has its roots in Social Learning Theory, which stipulates that aggression is developed in the context of learned social interactions and is further acquired through reinforcement (Vitaro & Brendgen, 2005). This type of aggression is intentional, and is exhibited when the individual aggresses for reasons other than social retaliation. Therefore, proactive aggression is intended so as to acquire perceived power, used in the domination of social groups or simply for the purpose of aggressing. Recent work by Baker, Tisak, Tisak, and Jensen (under review), arising from earlier versions of this study, suggest that it may also be a product of limited cognitive load, wherein a child is frustrated for non-social reasons and takes these frustrations out in the social domain.
Proactive aggression has been linked to home-life turbulence, such as increased levels of family violence, although not necessarily physical abuse of the child (Connor, Steingard, Cunningham, Anderson, & Melloni, 2004). Interestingly, proactively aggressive children tend to view family interactions more positively than do reactively-aggressive and nonaggressive children (Poulin & Dishion, 2000). One possible explanation is that these children have been raised with successful models of aggression in their home (Vitaro & Brendgen, 2005), such that children may learn effective behaviors from their parents or older siblings, and perhaps use this type of aggression as a tool, viewing it simply as a valence-free behavior. Children characteristic of this subtype have also reported having less parental monitoring and fewer rules in the home than do groups of nonaggressive children and reactively-aggressive children (Poulin & Dishion, 2000), which may suggest that parents of these children are more tolerant of such behaviors. However, this could also be a product of the increased likelihood of familial drug abuse (Connor et al., 2004), and in turn less parental monitoring. Of note, proactive aggression has been indirectly linked with later substance use (Connor et al.; Fite, Colder, Lochman, & Wells, 2007), and has been found to predict other delinquent behavior (Fite, Colder, Lochman, & Wells, 2008), whereas reactive aggression has not.

**Reactive Aggression.** Reactive or retaliatory aggression stems from the frustration-anger theory of aggression and therefore comes in response to another’s real or perceived provocation (Berkowitz, 1993; Dollard, Doob, Miller, Mowrer, & Sears, 1939, as cited in Vitaro & Brendgen, 2005), and thus is exhibited in anger when one’s goal is inaccessible, typically by means of another’s interference. The primary goal is to react to the provocateur and to cause him or her harm (Vitaro & Brendgen, 2005) or as a response to frustration. Therefore, this type of aggression is nearly immediate and impulsive.
Research on reactive aggression shows that social and environmental hostility and maltreatment are not uncommon among these children (Dodge & Coie, 1987; Dodge et al., 1997; Connor et al., 2004). These children often have poorly developed social information processing and tend to hold hostile attribution biases (Crick & Dodge, 1996; Dodge & Coie). Not surprisingly then, these children have difficulty with peer relations and with adjusting to peer expectations (Dodge, Lochman, Harnish, Bates, & Pettit, 1997). Generally, compared with their proactively aggressive cohorts, these children tend to be less accepted by peers (Prinstein & Cillessen, 2003; Vitaro & Brendgen, 2005), which may in fact impact the development of a hostile attribution bias (Crick & Dodge).

**Inhibitory Control and Aggression**

Inhibitory control has been linked to a number of social competency constructs, including emotion regulation (Carlson & Wang, 2007), internalizing behaviors (Rhoades, Greenberg, & Domitrovich, 2009), externalizing behaviors (Rankin, 2007; Utendale, & Hastings, 2011), and problem behaviors (Hughes, & Ensor, 2008). However, inhibitory control and aggression in particular have not been examined quite as thoroughly. The majority of research on this topic focuses on disorders, such as ADHD or ODD, rather than normative samples (e.g., Morgan & Lilienfeld, 2000; van Goozen et al., 2004). Other research on the topic is inconclusive; in a study of aggression and four aspects of executive functions (set shifting, inhibition, working memory, and verbal fluency), only inhibition correlated with aggression, in that aggressive children showed poorer inhibitory control skills compared to controls (Raaijmakers et al., 2008). In expanding the concept of effortful control further, both conduct behaviors and hyperactivity have been directly linked with delay of gratification, but only conduct problems were related with parent-reported inhibitory control (Gusdorf, Karreman, van Aken, Dekovic, & van Tuijl, 2011).
Prosocial Behavior

Prosocial behavior is understood as “any action that … benefits others, or promotes harmonious relations with others” (p. 33; Hay, 1994). In young children, this behavior is typically exhibited as sharing, cooperating, helping, et cetera, and is intentional (Carlo, 2006; Eisenberg et al., 2006). Evolutionary theory would suggest that prosocial behaviors are necessary for the betterment and success of the species, and that developing these abilities in early childhood is important for developing long-lasting relationships. That is, prosocial behaviors tend to develop relatively early in the lifespan, although there is not a crucial developmental time by which these behaviors must develop. Of course, there must be other factors that influence the development of prosocial behaviors, such as cognitive development and regular social interactions.

According to social cognitive theory, as children grow they develop the ability to perceive social cues and to internalize their meanings (Bandura, 1991). For instance, a child must first exhibit certain behaviors (positive or negative), or must view others as behaving in these ways. At this juncture, children are socialized by parents, siblings, and teachers via rewards and punishments. That is, children behave in positive ways, such as sharing or helping, and are praised (i.e., “What a good boy!”) or rewarded in other ways. Per theories on behaviorism, if this occurs frequently and consistently, children learn to associate exhibiting the behavior with the received praise, via positive reinforcement. Social cognitive theory, then, says that the child begins the process of internalization, in which the child begins to develop personal beliefs in line with the socialized behavior. These behaviors and their positive social meanings become of personal value to the child. By expanding their own schema of positive social interactions, children become more capable of successfully navigating social situations.
Theory of Mind and Prosociality

Advanced social behaviors, ToM, and moral understanding develop during the early childhood years, largely due to brain maturation and socialization, as previously discussed. Moral understanding, in which the child knows that another person is hurt or upset for example, typically develops more quickly than does ToM. Researchers believe that this occurs due to the abovementioned process of socialization; children are instructed from an early age that their behaviors have emotional consequences for both others (“when you share your toys with your brother it makes him happy!”) and themselves (“how would you feel if that happened to you?”), in line with both Bandura (1991) and Turiel’s (2006) theoretical viewpoints. Cognitive ToM, however, is less easily taught at such a young age.

Another vital component of the development of prosocial behavior is the understanding of empathy (Eisenberg & Liew, 2009). Empathy, one aspect of moral understanding, is considered to be the affective component of prosocial behavior in that people must be able to realize another’s point of view and further be emotionally compelled by that understanding (Lonigro, Laghi, Baiocco, & Baumgartner, 2014). This is considered in addition to the necessity of ToM in that empathy involves the internalization of the other’s affect, a component that is not necessarily a part of ToM, although still in line with social cognitive theory as explained above.

Social relationships in particular are heavily affected by ToM and moral development. For example, children with more developed ToM and moral perspectives have been rated as more social (Watson, Nixon, Wilson, & Capage, 1999), more socially competent (Lalonde & Chandler, 1995), as well as exhibiting more sophisticated pretend play (Leslie, 1987). Further, higher ToM scores have been linked to higher scores of morality (Boerger, 2015; Dunn et al.,
2000), and those with more developed ToM are considered more helpful and instructive with their peers (Strauss, Ziv, & Stein, 2002).

**Current Study**

In summary, recent literature points to a number of antecedents which predict social behaviors, namely aggression and prosociality. However, few studies explore all of the constructs examined here, particularly ToM and aggression simultaneously. For example, Olson, Lopez-Duran, Lunkenheimer, Chang, and Sameroff (2011) found that those with higher levels of aggression also exhibited lower levels of self-regulation and poorer ToM understanding. However, Olsen et al. operationalized aggression as a single construct, rather than being multi-faceted; furthermore, this study did not include morality or moral ToM. Previous versions of this study found links between ToM and IC, ToM and aggression sub-types, and IC and aggression (Baker et al., 2015), but did not examine morality. Finally, many studies examine ToM and morality, or morality and aggression, but few studies include all of these facets of socio-emotional development. Those that do examine all of these concepts, however, either used older samples (e.g., 8- to 14-years-old; Boerger, 2015), or focused more on peers’ perception of aggression, via peer acceptance, rather than on the child’s actual aggressive tendencies (Slaughter, Dennis, & Pritchard, 2002), which relies more on the child’s social judgments than others’ perceptions.

In the current study, I examined how developing socio-cognitive factors influence specific aggressive tendencies – choosing to aggress relationally versus physically, for example – as well as prosocial tendencies. The socio-cognitive factors of interest here are ToM, moral ToM, moral judgment, and lastly, IC. To do this, I utilized a battery of ToM measures, so as to offer a broad but in-depth understanding of this construct. Moral ToM was measured using a
single task as this is, to my knowledge, the only established measure of this emerging area of research. Moral understanding was measured using a singular task as well, as this measurement offers various opportunities for the child to express understanding. The last cognitive component in this study, IC, was measured by one task, resulting in two indices of IC: reaction time and accuracy. This follows theoretical and empirical (Baker et al., 2015; Hughes et al., 2000) support, in that reaction time should predict both reactive and proactive tendencies, and accuracy should be related with physical versus relational tendencies.

Behavioral outcomes were each measured using two paradigms. Recent research has shown that moral identity moderates the relation between ToM and aggression (Boerger, 2015), but not with prosocial behaviors, suggesting a more nuanced relationship. For this reason and as a follow-up to earlier research (Baker et al., 2015), aggression was operationalized as a multi-dimensional construct, encompassing both the expressive and cognitive components of aggression. Prosocial behaviors were examined using identical methodology as aggression, but was operationalized as a unidimensional construct.

**Hypotheses**

Specific aims of this study are as follows:

1. Theory of Mind (ToM) measures will be internally consistent such that the scores can be combined to create a single score for ToM. This will be true of the aggressive sub-types and prosocial scores as well; for example, the two measures for aggression will be reliable and congruent with one another so that a single score may be calculated for each type of aggression.

2. There will be a main effect of age: older children will all be more proficient on Inhibitory Control (IC), ToM, and Moral Theory of Mind (MoToM) tasks. Furthermore, age will
predict aggression, such that older children will show lower levels of physical aggression and higher levels of relational aggression, compared with younger children.

3. There will be a main effect of gender for physical aggression: males will be higher in physical aggression than females.

4. There will be a main effect of socioeconomic status (SES) for aggression; children of a lower SES will show higher levels of aggression than those of a higher SES.

5. IC scores will be positively related to ToM, such that those more proficient in IC will have higher ToM (see Figure 1).

6. Moral Judgment will predict MoToM – those who are more proficient on the Moral Judgment task will prove more successful on the MoToM task (see Figure 1).

7. MoToM and ToM will be positively related – those with higher MoToM scores will also have successful ToM expression (see Figure 1).

8. ToM will predict aggression, such that greater ToM will be associated with greater relational aggression, but not physical aggression (see Figure 1).

9. I anticipate interactions predicting each type of aggression:
   
a. **Physical-proactive aggression**: individuals that have developed ToM, but not MoToM, will still have higher levels of aggression; those with high levels of both ToM and MoToM should show lower levels of aggression.

b. **Relational-proactive aggression**: children high in ToM will show higher levels of relational-proactive aggression, except for those children that are also high in MoToM, who will show lower levels of this aggression.

c. **Physical-reactive aggression**: children high in ToM and IC will have lower levels of this type of aggression.
d. **Relational-reactive aggression**: a ToM-by-MoToM interaction will exist, whereby high levels of ToM precede higher levels of aggression, but only for children who are low in MoToM, and particularly for children from a low SES.
METHODS

Participants

One hundred seventy-six children ($M_{age} = 53.24$ months, $SD = 9.83$) were recruited from 13 public and private preschools in the Northwest Ohio region. The sample ranged in age between 33 and 78 months; there was 1 two-year-old (33 months), 56 three-year-olds ($M_{age} = 42.8$ months, $SD=0.39$), 73 four-year-olds ($M_{age} = 53.3$ months, $SD = 0.43$), 39 five-year-olds ($M_{age} = 64.74$ months, $SD = 0.59$), and 7 six-year-olds ($M_{age} = 74.86$ months, $SD = 0.60$; see Table 1 for all descriptive statistics). There were slightly more males (54.5%) than females (45.5%) in the sample, and the children were generally from affluent backgrounds ($M_{family\ income} = $116,093, $SD = $98,628, Median = $100,000; Range: [$19,200-900,000$]), and two-parent households (86.4%). Parents reported education levels for both parents (when available); mothers’ reported educational attainment was equal or equivalent to college graduate, with many indicating some graduate training (see Table 1), and the same was generally found for fathers’ education. Parents’ educational attainment tended to increase with the child’s age, such that parents of older children were more educated than parents of younger children (Mothers’ Education $r=.15$, $p<.05$; Father’s Education $r=.16$, $p<.05$).

Recruitment for this study involved first contacting area preschools via existing community relationships, cold-calling, and snowballing techniques. Either myself or two trained undergraduate research assistants, for whom I provided a script and training, contacted 37 preschools. Of the 37 schools, eight had been involved with previous studies and agreed to participate in this study; one school that had agreed to this study never completed follow-up, and a school that was contacted due to prior involvement in research was no longer in service, and so did not participate in this study. Two schools involved with previous studies had individually
agreed to this study but had actually combined into one school by the start of data collection, and so were considered as one school (School 2; see Table 2). Of the remaining preschools – those not involved with previous studies – 10 agreed to participate. However, although two of these schools agreed and sent home letters to parents, at the time the letters were returned to researchers the data collection for this study was completed and data analyses was already taking place, and so no data were collected for these sites and the school information was not included. Although no school actively responded that they would not like to participate, several schools did not respond to cold-calling and emails. Non-responsive schools were primarily affiliated with educational establishments (i.e., elementary schools), and so it is likely that they did not have the power to provide consent given the organizational structure.

Of the 13 schools that consented (see Table 2), six were from sub-rural communities, five were suburban, and two were in urban settings. They varied in regards to their funding models: five were non-profit organizations (N=45; 72.7% White, 11.4% African American, 11.4% Multi-racial; $M_{\text{family income}} = 80,000; SD = 39,703; Range: [$20,000-200,000]) which relied on federal funding and grants; four were private institutions that – although privately owned - were still controlled by community governance (N=54; 85.2% White, 5.6% Multi-racial; $M_{\text{family income}} = 98,015; SD = 52,878; Range: [$19,200-300,000]); three were independent schools (N=69; 72.5% White, 8.7% Multi-racial; $M_{\text{family income}} = 169,927; SD = 166,338; Range: [$31,500-900,000]), funded through tuition and alumni donations, and were independently governed; and one was religiously-affiliated (N=9; 100% White; $M_{\text{family income}} = 106,166; SD = 30,000; Range: [$77,000-160,000]) and funded through parishioner donation.

Once we received documented consent for the schools from either the director or individual teachers, parent permission packets were sent home to the parents from the schools.
These included a detailed letter describing the study and a form requesting basic demographic information. Two hundred parent forms were returned; of these, 18 parents did not consent to their children participating. Further procedural and pragmatic issues arose which prevented participation by 6 individuals for whom we had acquired parental consent: one did not speak fluent English, two were either sick or changed schools during interview times, one was not verbally capable of completing interviews, one had been referred to psychological services and so was excluded from interviews, and one was excluded due to lack of sufficient information from parents on the consent form. The excluded children – those for whom we had parental consent and those for whom parental consent was withheld – were varied in regards to their preschool’s socio-economic status, and so we believe that the exclusion of these children did not alter the findings of this study.

**Procedures**

Each child was individually interviewed three times: once for a standard verbal assessment – completed by a trained psychology graduate student – and twice for the primary interviews germane to the goals of the study – completed by one of four trained undergraduate research assistants. Interviews took place in a random order, and all interviews for a child took place within two weeks. Interviews took place at the child’s school, in a quiet area of the school or classroom selected by the teacher or director where there were no interruptions or distractions. Interview times were arranged by myself and the school director or teachers, and usually took place in the mid-morning or mid-afternoon.

The two primary interviews involved 5 ToM measures, 1 MoToM measure, 1 standard Morality measure, 1 IC measure, and the combined prosocial/aggression measure; as two of these tasks involve open-ended questions and one involves immediate response by both the
researcher and child, the primary interviews were all audio recorded. Audio files were uploaded to a secure server at the end of the interview day, and were coded with the participant number, interview number, and researcher’s initials. All audio files were coded, and transcribed (for open-ended questions) by two undergraduate research assistants, and inter-rater reliability was high for both categorical variables (Kappa=.91, p<.001) and continuous variables (such as reaction time on the Stroop task; r=.96, p<.001) on 20% of the cases (N=36 cases, or 72 interviews). Both undergraduate research coders were trained for appropriate procedures and coded per an instruction manual, which was created for this study and written such that responses (rather than correctness or accuracy) were recorded. Open-ended questions were first transcribed and then coded into categorical responses, per the coding manual designed for those tasks by Killen et al. (2011). Of note, while the open-ended questions were completed per the procedures, these questions were not included in the analyses, per the research proposal for this study.

Measures

Demographic Information

Basic demographic information, such as age, gender, ethnicity, socioeconomic status, and family makeup was collected by sending home to the parents a packet, including a description of the study, a consent form, and demographic questions.

Theory of Mind Battery

Five false-belief tasks were used to assess ToM (see Table 3). All measures have been found to be both valid and reliable, and have been used in prior research (e.g., Baron-Cohen et al., 1985; Hogrefe, Wimmer, & Perner, 1986; Perner, Leckam, & Wimmer, 1987; Wellman & Liu, 2004; Wimmer & Perner, 1983). The following ToM measures are described in order of increasing difficulty based on previously established Guttman scaling (Wellman & Liu, 2004);
past research shows the vast majority of preschoolers (95%) can pass the Diverse Desire task, but this is not true of the Knowledge Access task (73%), and even less so for the Appearance-Reality task (59%), the Explicit False Belief task (57%), and Real-Apparent Emotion task is the most difficult of all (32%; Wellman & Liu).

_Diverse Desires Task._ Designed in line with Wellman and Woolley (1990), this task assesses how children understand that another person can hold different desires than themselves. In this task, the child is told that a particular person prefers one snack (opposite of the child’s own taste) and dislikes another, and then the child must choose an appropriate snack for the other person. For this study, the child was presented with a toy figure, “Mr. Smith,” a picture of cake, and a picture of broccoli. The researcher prompted, “Here is Mr. Smith. It’s snack time, so Mr. Smith wants a snack to eat. Here are two different snacks: broccoli and cake. Which snack would you like best? Would you like a piece of cake or some broccoli?” (own-desire question; Wellman & Liu, 2004). If the child chose broccoli: “That’s a good choice, but Mr. Smith really likes cake. He doesn’t like broccoli. What he likes best is cake.” If the child preferred cake, then Mr. Smith liked broccoli and disliked cake. “So, now it’s time to eat. Mr. Smith can only choose one snack, just one. Which snack will you choose for Mr. Smith, the broccoli or the cake?” (target question). In order to be successful at this task, the child must have chosen the snack in the target question opposite than his answer for the own-desire question (Wellman & Liu). Success in this task was scored as 1, and failure as 0.

_Knowledge Access Task._ Similar to a paradigm originally designed by Hogrefe and colleagues (1986), this task is a widely used method of false-belief ToM assessment, utilized in a variety of versions. In this modified version (Pratt & Bryant, 1990; Wellman & Liu, 2004), a nondescript box was used in the following manner: the researcher presented the box to the child,
and asked the child, “What do you think is in this box?” (control question). The child could respond to this question in any way (e.g., by suggesting any type of object, or by saying “I don’t know”). Once the child answered the control question, the child was then instructed to open the box, revealing to the child that it held birthday candles. The box was then closed, and the child then presented with a toy figure of a child: “Davey has never seen inside this box. Now here comes Davey! So, does Davey know what’s inside the box?” (target question), “Did Davey see inside the box?” (control question). In order to be correct, the child must have answered the target question “no”, earning 1 point, else earning no points.

Appearance-Reality Task. First described by Flavell (1986), appearance-reality tasks attempt to help the child articulate the difference between what something appears to be and what it really is. In this version of this task, a sponge painted to look like a rock was presented to the child. The child was asked to identify the object, and then asked to pick it up and play with it, at which point the child was asked, “What is it, really and truly?” (control question). Following this, the child was asked what someone else would think it was if they walked into the room at that moment (target question). If the child indicated that another person would think it was a sponge, the child scored no points; if the child indicated that someone would think it was a rock, his/her response was coded as 1.

Explicit False-Belief Task. Designed by Baron-Cohen et al. (1985), the Sally-Anne paradigm is a standard measure of false-belief assessment. This study involved the researcher, the child, and used two puppets, a ball, a basket, and a box. The child was introduced to the two dolls: Sally and Anne. The following script was read to the child, with the researcher using the above items to illustrate the scene to the child.
“This is Sally, and this is Anne. I’m going to tell you a story about Sally and Anne, okay? Sally has a ball, and she puts her ball in the basket. Sally goes outside the room and can’t see in the room anymore. Anne comes in, and Anne picks up the ball from the basket and puts it in the box. Anne goes away. Sally comes back.”

The child was then asked, “What were the two dolls names?” (control question), and then, “Where do you think Sally is going to look for her ball first of all?” (target question). If the child’s answer indicates the basket, the response is scored as 1. If the child indicates the box, the response is scored as zero.

Real-Apparent Emotion Task. Stemming from initial design by Harris, Donnelly, Guz, and Pitt-Watson (1986) this task assessed a child’s understanding that one’s expressed emotion, smiling for instance, may not necessarily indicate his/her true emotion. This task began by asking children to identify the emotions expressed in three cartoon faces (happy, sad, neutral). The child was then shown a picture of a boy’s head from the back, so that the boy’s facial expression is not visible. The researcher read the child the following,

“This is a story about a boy. I’m going to ask you about how the boy really feels inside and how he looks on his face. He might really feel one way but look a different way on his face, or he might really feel the same way inside as he looks on his face. I want you to tell me how he really feels inside and how he looks on his face.”

If the child indicated understanding, the story continues:

“This is a story about Matt. Matt’s friends were playing together and telling jokes. One of the older children, James, called Matt a name and everyone
laughed. Everyone thought it was funny, but not Matt. But, Matt didn’t want the other children to see how he felt, because they would call him a baby. So Matt tried to hide how he felt.”

Three questions were used in assessing understanding of this story: “In the story, what would the other children do if they knew how Matt felt?” (control question); referring to the pictures of the three emotions, “How did Matt really feel, when everyone laughed? Did he feel happy, sad, or okay?” (target-feel question); and, lastly, “How did Matt try to look on his face, when everyone laughed? Did he look happy, sad, or okay?” (target-look question). In order to be successful at this task children must have indicated in the target-feel question that Matt would feel more negatively than he appeared in their answer for the target-look question. Possible solutions to this included “happy” for target-look, and “okay” or “sad” for target-feel; or “okay” for target-look and “sad” for target-feel. If children answered in either of the above ways, their responses were counted as 1; if not, 0.

For all of the Theory of Mind tasks, if a child did not pass the control questions, the question was asked again. After failing a second asking, that child’s data was not included in ToM analyses. Tasks were coded separately, but later combined to create a single score, ranging from 0 to 5. Again, higher scores on the ToM battery indicate greater ToM abilities.

Moral Theory of Mind (MoToM) Task

Following design and implementation by Killen et al. (2011), the MoToM protocol is similar to a standard ToM task but incorporates an accidental transgression by one of the characters. While this construct and measure is relatively new in the field, it was designed for a sample nearly identical to this (age range 3-7 years, middle-income sample). Furthermore, this
measure holds valid when compared with both prototypic moral transgression tasks and standard ToM tasks, integrating both concepts fluidly. Killen et al. (2011) showed that morally relevant ToM develops along a similar trajectory as standard ToM; that is, older children are more successful than younger children on both False Belief-Location and False-Belief-Content items. Additionally, Killen and her colleagues found that children tend to express false-belief understanding, both prototypic and moral, at similar rates; specifically, children are typically stable in passing both a standard ToM task and the MoToM task, or stable in failing both, rather than passing only one. When children did pass only one task, they were more likely to pass the standard ToM task than the morally embedded measure, which may suggest that children might find the MoToM task slightly more challenging.

The “Accidental Transgressor” vignette (Killen et al., 2011) is as follows:

“This is Tommy and this is Josh. Tommy has brought in a cupcake from home
and is keeping it in this paper bag. Tommy puts the paper bag on the table
then goes outside to play. Josh is cleaning up the classroom and sees the paper
bag. Josh throws the paper bag in the trash.”

Of note, this is a slight change from the original script by Killen and colleagues (2011): in the original script, the last sentence states, “Josh is helping the teacher clean up the classroom…” This was removed from the current phrasing due to the implication that Josh was aiding an authority figure, and previous research has shown this to alter children’s judgment of moral behavior (Goldstein & Tisak, 2006; Goldstein, Tisak, Perrson, & Boxer, 2006).

To aid in child understanding of the task, a series of pictures accompanied the script. Specifically, children were shown pictures of: character “Josh,” character “Tommy,” a cupcake, a bag, a trash can, and a table. The children were asked a series of follow-up questions to assess
belief, intention, act evaluation, actor evaluation, and victim emotion (Killen et al., 2011). For the Likert-type items, scores ranged from 1 (definitely not all right), to 4 (definitely all right; see Figure 2). The follow-up assessment items are as follows: (1) “What did Josh, the boy who threw out the bag, think was in the bag?” (child is shown pictures of cupcake and trash, but no further prompts are provided; false belief of the accidental transgressor); (2) “When Josh threw out the bag, did he think he was doing something that was alright or not alright?” (intention of the actor; Likert-type); (3) “Why?” (justification for actor intent); (4) “When Josh threw out the bag, do you think he was doing something that was alright or not alright?” (act evaluation; Likert-type); (5) “Why?” (justification for evaluation of act). The last three items referenced the victim’s actions: (6) “Now Tommy wants to eat the cupcake that he brought from home, where will Tommy look for his cupcake” (victim ToM); (7) “How will Tommy feel about losing his cupcake?” (attributions of the emotional state of the victim; Likert-type); (8) “How will Tommy feel about Josh?” (attributions of the victim emotions towards the accidental transgressor; Likert-type).

While this task assesses a number of constructs, including prototypic ToM, the primary item to be analyzed in measuring MoToM is item (2): “When Josh threw out the bag, did he think he was doing something that was alright or not alright?” (intention of actor; Likert-type). If the child responds that Josh thought he was doing something bad, he is indicating a lack of MoToM in that he cannot dissociate the harmful act (throwing out another’s cupcake) from the positive intent (helping). Therefore, scores for the MoToM item will be from 1 (really not all right) to 4 (definitely all right; see Figure 2), with higher scores indicating more developed MoToM.
Moral Transgression Task

A standard and established task of moral transgression understanding was used (Arsenio & Kramer, 1992; Killen et al., 2011). In this task, the child was read a script and shown a series of pictures depicting the characters and objects. The story read to the child was as follows:

“This is Diane, and this is Mary. Diane is playing on the swings outside. Mary comes over and pushes her off the swing so that she can get on it. Diane falls to the ground and hurts her knee.”

The accompanying pictures included: the character “Mary,” the character “Diane,” and a swing set. Six follow-up questions were asked: (1) “When Mary pushed Diane, did Mary think she was doing something that was all right or not all right?” (actor intent; Likert-type, see Figure 2); (2) “Why?” (justification for actor intent); (3) “When Mary pushed, do you think she was doing something was alright or not alright?” (act evaluation; Likert-type); (4) “Why?” (justification for act evaluation); (5) “How will Diane feel about getting pushed?” (attribution of victim’s emotions toward act; Likert-type); and (6) “How will Diane feel about Mary?” (attribution of victim’s emotions toward transgressor; Likert-type).

As with the MoToM paradigm, this task provides access to much information that might not be germane to the primary goals of this study. Therefore, the primary question in this task was (3) “When Mary pushed, do you think she was doing something was alright or not alright?” The Likert-type scale for this item was identical to the MoToM graded scale: (1) really not all right, to (4) definitely all right (see Figure 2).

Of interest for post-hoc analyses, coding of the justification questions for both the MoToM and the Moral Transgressions measures were in line with previous studies (Killen et al., 2011; Killen & Smetana, 1999), which organized justification responses into 6 categories: (1)
Inhibitory Control Task

Inhibitory control skills were measured using Gerstadt, Hong, and Diamond’s (1994) day-night Stroop task. This task has been widely used since its inception (e.g., Hughes, & Ensor, 2006; 2007; 2008; Montgomery, Anderson, & Uhl, 2008). In this task a training and pre-trial session occurred before 16 trials. In full, this measure took less than one minute for completion.

During the training session, the researcher presented the child with a card showing a yellow sun on a white background (*See Figure 3a*), and instructed the child, “When you see this card, I want you to say ‘night’” (Gerstadt et al., 1994). The child was then asked to repeat the word “night.” The researcher then removed the first card, and showed a card with a black background and with a moon and stars (*see Figure 3b*), and instructed the child, “When you see this card, I want you to say ‘day’,” and the child was asked to repeat “day.” During this task the child was trained to disregard the prepotent response (i.e., saying ‘night’ when presented with a moon, and ‘day’ when presented with the sun).

The researcher again showed the white sun card, with no instruction. If the child was hesitant, he or she was prompted with, “What do you say for this one?” If the child responded correctly, the child was praised, and then a trial session occurred with the black moon card. If the child successfully answered both of the trial questions, these questions were considered the first 2 questions of the testing phase, so as to prevent boredom (Gerstadt et al., 1994). However,
if the child did not correctly respond to both of these questions, the rules of the game were repeated for the child, and the trial questions were repeated as well.

In the testing phase of this task, there were 16 card-presentations, with the white sun card presented 8 times and the black moon card presented 8 times. Cards were presented in a pseudorandom sequence (that is, night (n), day (d), d, n, d, n, d, n, d, n, d, n, d, n, d, n, d) as designed by the Gerstadt and colleagues (p. 136, 1994). During the testing phase of this task, no feedback was given to the child. This task has proven nearly impossible for younger 3-year-olds, quite difficult for older 3-year-olds (i.e., mean age 3.5 years), and is seen to be progressively easier as children age with ceiling effects presenting at age 5 (Gerstadt et al., 1994).

This task was measured by: the number of correct responses, and reaction time for each question, which was converted to mean reaction time for each child. As with the ToM tasks, if a child responded correctly to each card, he received a score of 1; if incorrectly, 0, such that a higher score indicated a greater accuracy. Reaction time was measured as the time between when the card was presented to the child and when the child responded, and was calculated during coding using a stop watch.

**Aggression**

*Aggressive Behavior Questionnaire – Teacher Report.* In order to measure aggression, past research was used (Crick, Casas, & Mosher, 1997; Raine et al., 2006) to design a set of 16 questions, with four questions aimed at understanding each subtype of aggression (i.e., proactive-physical, proactive-relational, reactive-physical, reactive-relational; see Appendix A, Note), at which point existing data (Baker et al., 2015) was used in a principal component analysis to determine whether items fit with the sub-construct they were designed to measure. The primary component analysis revealed that the items loaded on the four components as predicted (all
primary loadings > .6); furthermore, this model explained 77.9% of the variance in teacher-reported aggression.

In turning these sub-scales into a questionnaire, the order of the questions were mixed, and provided to the child’s teacher (see Appendix A). These reported behaviors were assessed using Likert-style scoring in terms of behavior frequency, such that a child could be given a score from 0 (never) to 4 (much of the time) for each item. Mean scores were created for each child for each subtype of aggression, resulting in each child having four aggressive behavior scores from this questionnaire – each score between 0 and 4. Higher scores indicated an increased likelihood for that particular type of aggression to be present in the child’s behavior. Previous versions of this study show that these four subscales have good internal consistency and reliability (Cronbach’s α = .86, .82, .84, and .81, for Physical-Proactive, Physical-Reactive, Relational-Proactive, and Relational-Reactive, respectively; Baker et al., 2015).

Aggressive Behaviors – Child Interview. Using simplified versions of the items designed for previous versions of this study, as described above, a research assistant used two puppets to assess how each child believed he or she might respond in certain situations, similar in design to the Berkeley Puppet Interview (Kidwell et al., 2010; Measelle, Ablow, Cowan, & Cowan, 1998). For this interview, the child was introduced to two identical puppets, Iggy and Ziggy, who took turns making opposing statements about themselves (i.e., one indicating an aggressive behavior, the other not), and who then stopped to ask the child to describe himself/herself. For example, the item on the teacher-report, “This child will tear up another’s artwork for fun” was modified to be appropriate for first-person (puppet) expression for a preschooler: Iggy: “I like to rip up other kids stuff!”; Ziggy: “I don’t like to rip up other kids stuff!”; then together, the puppets would ask the child “How about you?” (see Appendix B). For the reactive aggression items, the
puppets were first told a situation (e.g., “what do you do when other kids won’t play with you?”), to make the items more understandable for children (see Appendix B).

This open-ended method of interviewing allowed children to respond in a manner they felt comfortable with, such as verbally (e.g., “I’m like Iggy”, “I don’t tell them that”), or nonverbally (e.g., pointing), which has proven reliable and correlates well with adult reports of this behavior. Previous studies show that children acclimate to the interview style quickly and respond more honestly during the open dialogue with the puppets than in typical interview techniques with adults.

Given the extensive amount of child-interviews in this study, rather than using all 4 subtype items for this particular procedure, only 2 items from each aggressive sub-scale were used for this child-interview (see Appendix B). The 2 items selected for each component of aggression were those that most strongly loaded on their respective dimension (all loadings > .65; Baker et al., 2015).

**Prosocial Behaviors**

*Child Behavior Scale – Teacher Report.* Ladd and Proﬁlet’s (1996) Child Behavior Scale – Prosocial Subscale was used to assess prosocial behaviors (see Appendix C). The scale consists of 8 items measured on a Likert-type scale, and had teachers rate behaviors in terms of frequency, ranging from 0 (never) to 4 (always). Scoring for this measure was modified from the original 3-point Likert (0 = not true, to 2 = often true) to the current 5-point Likert-type scale for consistency with the Aggressive Behaviors Questionnaire – Teacher Form. This measure has proven valid in comparison with observational teacher reports (*r* = .23, *p* < .01) during free play at school. While this measure was originally demonstrated with a slightly older sample (*M* = 5.5
years), these items were constructed from many earlier measures validated on preschool samples (e.g., Levine, Elzey, & Lewis, 1969; Walker & McConnell, 1988).

*Prosocial Behaviors – Child Interview.* To be consistent with the Aggressive Behaviors Interview, this session included the opposing puppets interview technique (Measelle et al., 1998), described previously. Again, this method has proven very reliable in comparison with adult reports, allows for an open-ended response from the child, and seems to elicit more honest and accurate responses from the children as they become engrossed in the puppets’ interactions. Items for this session were simplified versions (see Appendix D) of those used in the Child Behavior Scale – Prosocial Subscale (Ladd & Profilet, 1996), described above. Again, these items have shown good consistency when compared with observational reports from teachers ($r = .23, p < .01$).

**Verbal Skills**

*Peabody Picture Vocabulary Test – Version 4.* Verbal abilities, a hypothesized covariate, were assessed using the Peabody Picture Vocabulary Test – Version 4 (PPVT-4; Dunn & Dunn, 2007), in order to control for verbal skills. Past research indicates that verbal skills can be highly related to ToM development, and some research indicates that to measure ToM without a measure of verbal skills is improper. The PPVT-4 has shown to be a good measure of receptive vocabulary, and is useful in ensuring that the child is capable of understanding all of the measures. The PPVT - 4 uses a verbal- or nonverbal-, multiple-choice format. Scores correlate highly with measures of general intelligence, along with scores of reading and language (Dunn & Dunn). The PPVT - 4 contains up to 350 items, and takes approximately 15 minutes to complete with a preschool sample, and scores are obtained by subtracting from a ceiling score. Of note for this study, as the youngest child for whom we had parental consent was not yet 3-years-old, it
has been acceptable for use with children as young as 2.5 years of age. The PPVT - 4 shows good reliability and validity (Dunn & Dunn).
RESULTS

As we used a number of analytical techniques, the following section is arranged in the order in which hypotheses were presented, including proposed analyses and any supplementary analyses. First, this section begins by discussing the scale reliability and construction for three scaled constructs: ToM, Aggression, and Prosociality, which were the focus of Hypothesis 1. The second set of hypotheses – hypotheses 2, 3, and 4 – were all aimed to address demographic variables (age, gender, and socio-economic status, respectively); the third group of hypotheses (hypotheses 5, 6, 7, 8) primarily examined relations between constructs of interest. The final hypothesis, and its sub-hypotheses (9a, 9b, 9c, 9d), addressed moderating relationships that predict specific social outcome behaviors.

Hypothesis 1: Scale Reliability and Construction

Hypotheses 1. Theory of Mind (ToM) measures will be internally consistent such that the scores can be combined to create a single score for ToM. This will be true of the aggressive sub-types and prosocial scores as well.

Theory of Mind

First, any children who did not answer any of the control questions were excluded to maintain validity. This revealed a surprising finding in that only 6.3% of children passed the control question for the Appearance-Reality task, compared with more than 82% for control questions on all other tasks (Diverse Desires, 100%; Knowledge Access, 83.5%; Explicit False-Belief, 98.3%; Real-Apparent Emotion; 82.4%). This finding likely indicates either measurement or researcher error, and so this task was dropped from further analyses. Once this occurred, the final sample of children who passed all ToM controls was 121. The restricted sample did not significantly differ from the full sample in regard to age distribution (Full Sample $M_{age} = 53.24$,}
SD=9.83; Restricted Sample $M_{age} = 54.34$, SD=9.68, $p=.21$). Of interest to other analyses, I tested whether or not the restricted sample differed in regard to socio-economic status and gender; the two samples were not significantly different in regard to income ($t(120)=-0.40$, $p=.69$), nor were they different in regards to proportion of females in the samples (see Table 5).

Next, I wanted to ascertain whether or not pass-fail rates on specific tasks were in line with findings from previous studies and in line with the expected trajectories. Passage rates for the current study are as follows: Diverse Desires, 80.2% of the sample answered correctly; Knowledge Access, 68.6%; Explicit False Belief, 57%; and Real-Apparent Emotion, 34.7%. These values are in line with previous studies (Wellman & Liu, 2004) in both passage rates and general progression of difficulty – Diverse Desires as easiest, Real-Apparent Emotion as most difficult.

In an effort to maintain a 5-item ToM scale, I investigated whether or not the ToM portion of the MoToM (“Josh-Tommy”) task could be used to replace this item, as this portion of the MoToM was originally designed to capture traditional ToM simultaneously with MoToM (Killen et al, 2011). The question, “What did Josh, the boy who threw out the bag, think was in the bag?” was used by Killen et al. to measure validity with ToM, and was intended for that purpose in the current study as well, as the question clearly measures False-Belief understanding and is very similar to the target question in the Explicit False-Belief task in this study (“Where do you think Sally is going to look for her ball first of all?”). Success on these two items was significantly related ($r=.23$, $p<.05$). Furthermore, although success on the morally imbedded ToM task was not related to the sum of the four ToM tasks using traditional significance testing

---

1 Because the unrestricted sample and ToM-controlled restricted sample are of different sizes and are not independent, I conducted a single sample $t$-test whereby the unrestricted sample mean(s) was used as the population mean to which the restricted sample means(s) was compared.
standards \((r=.15, p<.10)\), this is not unexpected whatsoever; these tasks were selected so as to be varied in difficulty (Wellman & Liu, 2004) and the morally imbedded ToM task was known to be the most difficult (Killen et al., 2011), and so scores on this one item would only be related to ToM battery scores for a small group of children (those with the most developed ToM), but would seem unrelated for others. Furthermore, this value is consistent with previous work on ToM test batteries, such as McAlister and Peterson (2007), which also highlight that a correlation coefficient is not the best method of measuring the inherently scaled nature of ToM batteries. Therefore, based on the previous research discussed above, the calculated value provided support to the hypothesis that this item would be an adequate replacement and could be used in the aggregate ToM score.

In order to create an aggregate score for each individual, using each of the 5 ToM tasks (4 original and 1 supplemental), I first tested for reliability. To do this, I used Rasch modeling (one parameter logistic; 1PL) in IRTPRO (version 2.1; Cai, Thissen, & du Toit, 2011); this analysis assumes that tasks progress in difficulty and, therefore, that there is a predictable trajectory of task success throughout the sample (Rasch, 1993). As a function of using multiple tasks that attempt to measure various levels of development, standard internal reliability coefficients, such as Cronbach’s alpha, would be inappropriate as this assumes that responses on one item should necessarily be in line with responses on all other items with the same focus. That is – using the current example of ToM – if a person passed one task, standard reliability tests would assume that the person would pass the other tasks as well. A Rasch model, instead, assumes that there are gradations of difficulty, and that if a person passed a particular task, then that person probably passed the easier tasks as well, but not the more difficult tasks. Furthermore, Rasch models have been used to assess the reliability of ToM aggregate scales in previous studies (e.g., Wellman &
Liu, 2004) and are particularly useful in developmental studies (Weiland & Yoshikawa, 2013). Lastly, the Rasch model is designed for use with items for which responses are dichotomized, such as with the current pass/fail items.

The results from the Rasch model showed that the 5-item ToM scale was an acceptable fit to the data (RMSEA = .09; see Table 6) and indicates that the selected tasks adequately represent interval scaling, in line with previous research. That is, the results replicate previous findings (Wellman & Liu, 2004) using both these specific tasks (with the exception of the supplemental item) and Rasch modeling, such that each of the tasks were of appropriate difficulty levels for a non-clinical, preschool sample. Furthermore, the model fit indicates that this sample tended to respond in a predictable progression – meaning that if a child passed the Explicit False Belief task, for example, he or she also probably passed the Diverse Desires task and Knowledge Access task, but not the Real-Apparent Emotion or supplemental False Belief task – as expected. Therefore, I feel confident that the five ToM tasks are consistent and reliable, and so for all further analyses regarding ToM the aggregate score was used.

**Aggression**

*Teacher Report.* The four proposed subtypes – physical-proactive, physical-reactive, relational-proactive, relational-reactive – were created per item specification found in Appendix A, which used the teacher-report data, and subtypes scores were calculated by averaging the responses for the four items. Although each subtype scale was internally consistent, the correlations between some of the subscales were sufficiently high in multiple cases ($r>.80$) as to merit considering the adoption of a two-factor model of aggression rather than the proposed four-factor model. The resulting two-factor model allowed for examination of physical and relational aggression. The reliability for each subscale remained high ($\alpha = .93, .91$; Physical and Relational,
respectively), and the two subscales were strongly related, as expected \((r=.77, p<.01;\) see Table 7). Teachers generally reported infrequent aggressive behaviors (Physical: \(M = 0.60(0.74)\); Relational: \(M=0.64(0.69)\); Range for both: [0-4]).

**Child Report.** To align with the two-factor model gained from teacher-reports, child-reported aggressive behaviors were organized into two factors as well (physical and relational). Each scale was originally made up of four items, but one item was dropped from each due to lack of inter-item correlation. Of note, the two items that were dropped were not similar in terms of content, phrasing, response order, or aggressive sub-type; one item was originally physical-reactive aggression, and the other was relational-proactive.

The resulting 3-item scales both showed lower reliability than the teacher reports (\(\alpha = .60, .33;\) Physical and Relational, respectively); however, these values are being used for several reasons: (1) this was a measure developed for this study, and so lower values of reliability are to be expected (2) the number of items in a scale is directly related to Cronbach’s alpha, and so the reliability may be lower due to fewer items, and (3) previous studies examining similar behaviors during early and middle childhood using child self-report data report equivalent or lower values of reliability (Edelbrock, Costello, Dulcan, Kalas, & Conover, 1985), and have shown that self-report reliability is directly positively related to age. This last point highlights the inherent difficulty in working with children of this age, and – when combined with the second point – provides support as to the usefulness of the current conceptualizations given the reliabilities.

The two subscales of child self-reported aggression were moderately correlated to one another \((r=.32;\) see Table 7), as expected. Furthermore, there was only one correlation among teacher-reported and child self-reported aggression, between teacher-reported physical aggression and child self-report of relational aggression. This relationship was weak \((r=.20)\), but
significant. The lack of further relations between report sources mirrors previous findings of child self-report and adult report (Edelbrock, et al., 1985). Children reported low levels of both physical ($M=0.33, SD=0.67$) and relational ($M=0.36, SD=0.57$; Range for both: [0-3]) aggressive behaviors.

**Prosocial Behaviors**

*Teacher Report.* The 8-item prosocial subscale developed for this study showed acceptable reliability ($\alpha = .92$), and was negatively related to teacher-reported physical ($r=-.64$) and relational aggression ($r=-.49$; see Table 7), indicating high convergent validity. Generally, teachers reported children as behaving prosocially much of the time ($M=2.97(0.72)$; Range: [0-4]).

*Child Report.* Children’s prosocial behavior was assessed by 8 items designed to mirror those used for the teacher-reports. Given that this subscale had nearly three times the number of items than did the child self-reports of aggression, the reliability was considerably higher, and was acceptable ($\alpha = .77$; see Table 7). Children gave moderate endorsement for prosocial behaviors ($M=1.95, SD=0.75$; Range: [0-3]). Child self-report of prosocial behavior was negatively related to child self-report of relational aggression ($r=-.20$) and teacher-report of physical aggression ($r=-.19$), although both of these relations were weak.

**Hypothesis 2: Effects of Age**

*Hypothesis 2: There will be a main effect of age: older children will be more proficient on Inhibitory Control (IC), ToM, and Moral Theory of Mind (MoToM) tasks. Furthermore, age will predict aggression, such that older children will show lower levels of physical aggression and higher levels of relational aggression, compared with younger children.*
A series of regressions were conducted such that the cognitive and behavioral constructs—both IC indices, ToM, MoToM, and physical and relational aggression—were individually regressed onto age. That is, age was entered as a predictor variable in each of the separate regression analyses, in which the cognitive and behavioral constructs were individually examined as criteria.

Age did not predict IC-accuracy (see Table 8 for all coefficients and values for explained variance). Reaction time on the Stroop task was significantly predicted by age, such that older children responded more quickly than did younger children. ToM was predicted by age, such that older children were more successful on the ToM battery. Physical aggression—reported by either teacher or child self-report—was not predicted by age. Relational aggression was not predicted by age by either child self- or teacher-report. Age did predict both teacher- and child self-reports of prosocial behavior—older children were more prosocial (see Table 8).

Contrary to previous findings (Killen et al., 2011), age was not related to children’s understanding of the target character’s motive in the MoToM story, which was used to assess MoToM. This finding does not support hypothesis 2; it is important to note, however, that Killen et al. (2011) found age differences between only their youngest group (which they called “3.5 year olds”, $M_{age}=49$ months) and their oldest group (7.5 years, $M_{age}=92$ months), and no other age differences were reported. That is, the original study did not explicitly examine MoToM development in younger preschool-age children, and found no differences for the other age groups paralleled in the current study. Therefore, it may be possible that ages of interest in the current study do not differ in their conceptualizations of morally-relevant mental representations, as this is a more elegant and advanced cognitive ability compared with ToM, which may develop during the middle- and late-preschool years. In comparing the mean level of response by age,
results indicated that younger children tended to respond with greater understanding than did the older groups, with the exception of the oldest group of children; that is, 3-year-olds indicated greater understanding than did 4- and 5-year-olds, which is developmentally inappropriate and somewhat confusing.

To help explain this finding, I explored responses concerning the *justifications* for their answers. Frequencies of responses, separated by age, revealed that 65% of 3-year-olds gave a justification that lacked substantive rationale. Specifically, 32% of 3-year-olds failed to provide any response; 13% gave a response that was uncodeable (e.g., “because”; “I don’t know”); and 19% responded with an undifferentiated “good” or “bad”. Comparatively, 4-year-olds most frequently responded that the character thought he was doing something bad because he was damaging property (e.g., “He threw away his cupcake”) or they failed to respond entirely; 5-year-olds most frequently responded that the character thought he was doing something wrong due to property damage or psychological harm (e.g., “He’ll be sad”); and 6-year-olds most frequently indicated that it was *somewhat okay* or *okay* due to lack of negative intent (e.g., “he didn’t mean to do it”; “he didn’t know the cupcake was in there”). In summary, as children got older they tended to provide more elaborate and developed justifications as to why the behavior was/was not acceptable.

**Hypothesis 3: Effects of Gender**

*Hypothesis 3: There will be a main effect of gender for physical aggression: males will be higher in physical aggression than females.*

Two paired samples t-tests were conducted, such that mean levels were compared for physical aggression (1) as rated by teachers, and (2) per the child self-report, and responses were separated by gender. Results for both were not significant; males and females did not differ in
their expression of physical aggression according to teachers’ perception or according to their own beliefs. Gender did not have any effect on perceived levels of aggression, and was therefore not included in further analyses (see Table 9).

Although I had not hypothesized main effects of gender on the other variables, hierarchical regressions were used to test for effects of age and any age-by-gender interactions on all of the variables of interest: IC-accuracy, IC-reaction time, ToM, MJ, MoToM, and teacher and child self-reports of physical and relational aggression, as well as teacher and child self-reports of prosocial behavior. No effects of gender were found for any variables, nor was any age-by-gender interaction revealed for any variable.

**Hypothesis 4: Effect of Socio-economic status on Aggression**

*Hypothesis 4: There will be a main effect of socioeconomic status (SES) for aggression; children of a lower SES will show higher levels of aggression than those of a higher SES.*

A series of regression analyses were used in which the four reports of aggression – Teacher reports of physical, relational; Child self-report of physical, relational – were individually regressed upon family income (reported by parents, and used as a proxy for socioeconomic status). Physical aggression, per teacher report, was not predicted by the socioeconomic status of the child’s family nor did the socioeconomic status predict the child’s perception of his/her own aggression. Therefore, the socioeconomic status of the child’s family did not have any impact on the perceived aggressive tendencies of the child and was not controlled for in further analyses.

**Hypotheses 5 – 8: Relations among Primary Constructs of Interest**

*Hypothesis 5: IC scores will be positively related to ToM, such that those more proficient in IC will have higher ToM (see Figure 1).*
Hypothesis 6: Moral Judgment will predict MoToM – those who are more proficient on the Moral Judgment task will prove more successful on the MoToM task (see Figure 1).

Hypothesis 7: MoToM and ToM will be positively related – those with higher MoToM scores will also have successful ToM expression (see Figure 1).

Hypothesis 8: ToM will predict aggression, such that greater ToM will be associated with greater relational aggression, but not physical aggression (see Figure 1).

In order to examine relationships among primary variables – IC-accuracy, IC-reaction time, ToM, MoToM, MJ, physical aggression, relational aggression, and prosocial behavior – a set of path analyses were constructed in AMOS (Version 22.0.0) for SPSS, such that all hypothesized relations were accounted for (see Figures 4 & 5). The first set of path analyses used the teacher reports (see Figure 4) of outcome variables (physical and relational aggression, and prosocial behavior), and the second set examined the child self-report (see Figure 5) of outcome variables; the two sets were otherwise identical. Both sets of analyses, and any re-specifications thereafter, were completed using the default maximum likelihood method of estimation, which allows for estimation in the case of missing data using the full information maximum likelihood method. Of note, teacher-reports were missing for one child in the ToM-control sample due to teacher illness; this missingness was interpreted as random, and therefore unlikely to influence results. As many children did not pass the control questions for the ToM battery, as previously discussed, the following models only included children who passed each of the 5 ToM-control items (N=121).2

---

2 As a general rule, most SEM models require N=200, but there is great variability in this especially in regard to path analysis (Kline, 2016); to assess for issues of sample size, I ran a power analysis (MacCallum, Browne, & Sugawara, 1996; Preacher & Coffman, 2006) which indicated that the minimum sample size needed to detect effects was 88. This number is less than
Path Analyses: Teacher Report

*Primary Sample.* The proposed model (see Figure 4) – aimed at examining teachers’ perception of aggressive and prosocial behaviors – was not a good fit to the data, $X^2(24, N=121) = 64.65, p < .05$. Additional fit indices were likewise not acceptable: CFI = 0.92 (values greater than .95 indicate good fit; Hooper, Coughlan, & Mullen, 2008); RMSEA = 0.12 (values less than .08 indicate acceptable fit; Hooper, Coughlan, & Mullen, 2008). The 90% confidence interval RMSEA values [0.08, 0.15] does not allow for the acceptance of the close-fit hypothesis and for the rejection of the poor-fit hypothesis (Kline, 2016). Therefore, the proposed model was an inadequate representation of the data using teacher reports gained during this study.

Recall, however, that this sample included some individuals who seemed to have poor understanding of the MoToM task – there was a weak negative relationship between age and MoToM, such that younger children reported greater acceptance of unintentionally harmful behaviors, which would have indicated greater ToM than is developmentally plausible for their age. Therefore it is likely that the scores for this measure are not useful when considering the full sample of children. That is, when considering the full range of ages (3- to 6-years-old), the MoToM variable is not useful due to the scores earned by the 3-year-olds, and it is therefore quite likely that this misrepresented construct was affecting the fit of the model. To account for this, multiple sample compositions were attempted to correct these scores.

*Sample Composition 1: MoToM Initial Restriction.* In searching for more appropriate considerations of the MoToM scale, the original sample was restricted to include any children whose answer to “When Josh threw out the bag, did he think he was doing something alright or not alright” (dichotomous) was in line with their answer for, “How alright/not alright was it?”
(scaled; see Figure 2). For example, if they responded “not alright” to the first question but answered 3 or 4 (indicating that it was either alright or really alright) to the second question, this indicated poor task understanding and these children were thusly excluded. This restriction allowed for a sample of 109 children. This restriction was not a good fit to the proposed model – $X^2(24, N=109) = 43.59, p < .05; \text{CFI} = 0.95; \text{RMSEA} = 0.08 \ [0.04, 0.12]$ – although appeared better in comparison to the previous subsample.

**Sample Composition 2: MoToM Conservative Restriction.** In the second attempt to correct for the erred measurement of MoToM, I considered the justifications provided by the children as to why they supposed that Josh believed he was doing something alright/not alright, as reiterated in the previous paragraph. Again, these justifications were coded according to Killen et al.’s (2012) original conceptualization, but of interest for this current restriction were those whose responses were uncodable or otherwise unrelated to the task. This restriction - which was made in addition to the previous restriction - was rather conservative in regards to cognitive development; for example, young children may be able to acknowledge that something is “not alright” but may not be able to articulate their justification for this, which should not imply that they cannot discern between acceptable and unacceptable behavior. However, given the measurement issue, this approach is a logical extention to the previous restriction. The resulting sample included 84 children, and the model was of acceptable fit: $X^2(24, N=84) = 34.24, p = .08; \text{CFI} = 0.97; \text{RMSEA} = 0.07 \ [0.00, 0.12]$ (see Figure 4).

**Sample Composition 3: Age Restriction.** The last alternative sample composition was done by restricting the sample based on age, such that any children under 3.5 years were excluded. As previously stated, the three-year-old group seemed to struggle with the MoToM in comparison with the other groups, and so this was done to see if there were differences based on
age, rather than cognitive development, as explored with the previous compositions. This restriction resulted in including 110 children, and was not a good fit to the proposed model ($X^2(24, N=110) = 59.53, p < .05$; $CFI = 0.92$; $RMSEA = 0.12$ $[0.08, 0.15]$).

In summary, the more conservative composition was acceptable. Generally, this indicates that the conceptual model is an adequate representation of the theoretical relations between social cognitive factors and child behaviors. Of interest to the goals of the current study, measurements of social cognition significantly predicted prosocial behaviors (see Figure 4). That is, both ToM and MoToM significantly predicted teacher-reports of children’s prosocial behaviors, such that with better-developed ToM, and greater understanding of the target character’s motive in the MoToM story, reports of the child’s prosocial behavior tended to be more positive. Furthermore, estimates indicated an interaction between ToM and MoToM on prosocial behavior. In order to understand this further, post-hoc analyses were run using PROCESS for SPSS (discussed below).

**Path Analyses: Child Self-Report**

*Primary Sample.* The model used to assess the adequacy of the child self-report data (see Figure 5) – which was identical to the primary model of the teacher report, with the exception of the source of outcome variables – was not a good fit: $X^2(24, N=121) = 67.69, p < .05$. Additional fit indices were likewise poor ($CFI = 0.88$; $RMSEA = 0.12$; $90\% CI [0.09, 0.16]$), and did not provide support that this model fit the unrestricted sample using child self-report.

*Sample Composition 1: MoToM Initial Restriction.* In using the more liberal restriction of MoToM indicators, the model was not a good fit to the restricted data: $X^2(24, N=109) = 62.93, p < .05$; $CFI = 0.87$; $RMSEA = 0.12$ $[0.09, 0.16]$. 
Sample Composition 2: MoToM Conservative Restriction. The most restrictive composition of the sample was also not a good fit to the proposed model: \( X^2(26, N=84) = 49.48, p < .05; \text{CFI}=0.91; \text{RMSEA} = 0.11; 90\% \text{ CI}[0.07, 0.16] \). These values indicate that this sample specification - when using child self-report – did not fit the hypothesized model.

Sub-sample 3: Age Restriction. Using the age-restricted sample composition, the model was not a good fit to the data, \( X^2(24, N=110) = 70.32, p < .05; \text{CFI} = 0.86; \text{RMSEA} = 0.13 [0.10, 0.17] \).

To sum, when using the teacher reports of social behavior, the proposed model was acceptable when the sample was strictly constrained to include only those capable of reflecting their understanding of the MoToM task. As this sample restriction is the most theoretically justified way to account for the erred measurement with the younger children, only findings from the conservatively restricted sample will be discussed further.

In terms of support for specific hypotheses, Hypothesis 5 was not supported when accounting for all variables: neither IC-accuracy nor IC-response time significantly predicted ToM in the restricted sample. Likewise, Hypothesis 6 was not supported when examined within the scope of path analysis: Moral Judgment did not predict MoToM in the restricted sample. Hypothesis 7 supposed that ToM and MoToM were related; although not present in Figures 4 and 5, the error terms for these variables were not related when accounting for all theorized relations. Lastly, Hypothesis 8 was not supported using path analysis: ToM did not predict physical or relational aggression; however, physical aggression was predicted by MoToM when using child self-report, such that greater scores on MoToM predicted greater aggression. In sum, while the majority of these relations were individually non-significant, the model was still a good fit to the data using teachers’ perceptions of behaviors. This could indicate that the analysis was
underpowered due the restricted sample size. That is, although initial power analyses indicated sufficient sample size, a post-hoc sensitivity analysis (i.e., G*Power, 3.1; Faul, Erdfelder, Buchner, & Lang, 2009) indicated that smaller effect sizes might not be detectable using the current model and sample size. Specifically, the proposed model – using the conservative sample restriction – was powerful enough to detect small to moderate effect sizes ($r = .26$) in individual relations.

Given that none of the hypotheses were supported per individual path estimates – conflicting with much prior research – follow-up hierarchal regression analyses were conducted to test individual relations. This was to account for the lack of power in the path analysis, which can be due to the number of predictor variables in the model. For each of the following analyses, age – and verbal skills when ToM was considered an outcome variable – were entered in Step 1, and the predictor was entered at Step 2. Regarding hypothesis 5 – predicting ToM from IC – regression results were non-significant for both accuracy, $t(77) = -0.001, p > .05$, and response time, $t(77) = -0.03, p > .05$, when controlling for both age and verbal skills. Hypothesis 6 was supported: Moral Judgment significantly predicted MoToM - $t(81) = 2.00, p < .05, \Delta R^2 = .05$ – when controlling for age. MoToM and ToM were not related when controlling for age and verbal abilities ($r = .01, p > .05$), disconfirming hypothesis 7. In summary, moral judgment significantly predicted understanding of another’s motives, such that the better the child could identify moral/immoral behavior, the better the child was at identifying whether or not a person’s unintentionally harmful actions were acceptable from that person’s perspective.

Regarding hypothesis 8 – that greater ToM would predict greater relational aggression – this was not supported using either teacher- ($t(79) = -0.44, p > .05$) or child self-report ($t(81) = -0.14, p > .05$) controlling for age; physical aggression was, however, negatively predicted by ToM
using child self-report controlling for age ($t(81)=-2.32, p<.05$); this was not found for teacher-reports of physical aggression ($t(79)=-0.26, p>.05$) when controlling for age. Furthermore, understanding of another’s motive (MoToM) did not predict physical ($t(79)=-0.69, p>.05$) or relational aggression ($t(79)=-0.16, p>.05$) using teacher-report, nor child-reported relational aggression ($t(81)=0.84, p>.05$) when controlling for age, but it did significantly predict child self-reported physical aggression - $t(81)=2.28, p<.05$. In sum, child self-report of physical aggression was predicted by both ToM and MoToM, but in opposite directions: greater mental state understanding predicted lower levels of aggression, while greater understanding of other’s motives predicted greater levels of aggression.

**Hypothesis 9 - ToM, MoToM, and Behavior: Moderation Analyses**

_Hypothesis 9: I anticipate interactions predicting each type of aggression:_

- **a. Physical-proactive aggression:** individuals that have developed ToM, but not MoToM, will still have higher levels of aggression; those with high levels of both ToM and MoToM should show lower levels of aggression.

- **b. Relational-proactive aggression:** children high in ToM will show higher levels of relational-proactive aggression, except for those children that are also high in MoToM, who will show lower levels of this aggression.

- **c. Physical-reactive aggression:** children high in ToM and IC will have lower levels of this type of aggression.

- **d. Relational-reactive aggression:** a ToM-by-MoToM interaction will exist, whereby high levels of ToM precede higher levels of aggression, but only for children who are low in MoToM, and particularly for children from a low SES.
Hypothesis 9 – as outlined above – was not analyzable due to issues of measurement of aggressive sub-types within this sample, as previously discussed (see Results – Hypothesis 1: Aggression). For the current study, moderation by MoToM was explored for the three behavioral outcome variables: prosocial, physical aggression, and relational aggression.

The use of path analyses was necessary to understanding the complexity of the hypothesized relations between all variables; given the presence of an interaction in predicting prosocial behavior using teacher-reports, follow-up analyses were conducted using PROCESS (Hayes, 2013) for SPSS (Version 2.15), specifically utilizing pre-constructed Model 1. As PROCESS does not allow for moderation within serial mediation, ToM was entered as the independent variable, MoToM was entered as the moderator, and teacher-reported prosocial behavior was entered as dependent; all other variables in the model were entered as control variables.

Results of the test of moderation indicate that – although there was no main effect for mental state understanding nor understanding of motive on prosocial behavior – there was a ToM-by-MoToM interaction, such that for children who did not indicate having understanding of motive, there was no relationship between mental state understanding and prosocial behaviors; however, children who understood motive tended to express less prosocial behavior, per teacher report, as they understood others’ mental states (see Figure 6). This interaction explained 8% of the variance to the overall model in explaining prosocial behavior – $F(1,71)=6.24, p<.05$.

A test of MoToM moderation was also conducted for the other outcome variables: teacher-reported physical and relational aggression, and child self-reports of prosocial behaviors and physical and relational aggression. Results indicate that within the current study there existed no moderation by MoToM on any other behavioral outcomes.
DISCUSSION

Much research in recent years has focused on the relations between ToM and social competence, as well as ToM understanding and moral judgment. The current study suggested that these two areas are in fact interrelated with regard to early child development, and therefore contributes to this growing field by specifically examining children’s and teachers’ perceptions of social behavior as outcomes of socio-moral cognition. By extending theoretical framing by Bandura (1989) and Turiel (1983), I was able to provide some evidence that young children’s social behavior may be influenced by their understanding and application of other’s mental states with regards to moral intention.

A body of research has developed which suggests that more advanced mental state understanding is necessary for the application of moral cognition; understanding a person’s motive is considered a higher-order and applied form of moral understanding adjacent to ToM development, whereby an action can still cause harm although the actor is not to blame due to lack of intent. Relatively few studies have tested this hypothesis directly (Boerger, 2015), as this is still a new area of study, and even a smaller number have examined this within the scope of early childhood. While previous studies have found that older children are quite adept at discerning action harm and actor intent, younger children often assume negative intent as a function of negative outcomes (i.e., if something bad happens, this person was by necessity being bad; Margoni & Surian, 2016). Therefore, with age children seem better able to use mental state understanding to view motive as separate from outcome valence (Cushman, Sheketoff, Wharton, & Carey, 2013).
Review of Findings and Contributions

First, much of the findings from the current study – with the exception of the MoToM moderation, for instance – mirrors previous empirical work in the field, and therefore will be discussed briefly in terms of past research. After this, the new contributions of the current study will be discussed at length, as these are novel findings and therefore are of greater interest.

Review of Individual Relations

The current study found that ToM abilities follow a predictable trajectory, and that this trajectory can be measured using existing paradigms, in line with previous work by Wellman and Liu (2004). In regards to measuring aggression and prosocial abilities, although the teacher survey of aggression did not yield data providing for sub-type aggressive measurement, the resulting two-factor structure is still in line with myriad previous work (Crick & Grotpeter, 1995; Vitaro, Brendgen, & Barker, 2006) and supports the consideration of aggression wherein physical and relational aggression are conceptually different. Child self-report of aggression did not yield satisfactory reliability scores per traditional standards, although this is not surprising for self-report for children this young (Edelbrock, Costello, Dulcan, Kalas, & Conover, 1985).

Age, in line with much prior research, was a predictor of inhibitory control skills (Sabbagh, 2006) and ToM (Laranjo, Bernier, Meins, & Carlson, 2010; Siegal & Varley, 2002). Age did not strongly predict relational or physical aggression, which is not surprising giving the age of the current sample—in early childhood relational aggression is still developing and therefore this function may not be a fully operational mode of behavior, which could make this behavior inconsistent and perhaps infrequent. Per the developmental theory of aggression (Bjoerkqvist, Lagerspetz, & Kaukiainen, 1992; Bjoerkqvist, Oesterman, & Kaukiainen, 1992), inconsistent relational aggression would imply inconsistent use of physical aggression as well.
Furthermore, age did not predict MoToM scores, although further examination showed that the youngest children may not have understood this task as well as expected—this issue will be discussed in full later. Lastly, there were no effects of gender; while this finding is not surprising (see Ambrose & Menna, 2013; Salmivalli & Kaukiainen, 2004; Valles & Knutson, 2008), this may again be due to the age of the sample, in that gender roles may not be as internalized at such a young age, or the behavioral regulation process may be inconsistent.

In regards to the relations between primary constructs, although individual path estimates in the path analysis were non-significant (as discussed next), moral judgment did predict applied moral cognition (MoToM) when controlling for age—greater understanding of wrongness predicted greater understanding of another’s intent, in support of original work by Killen and colleagues (2011). Although ToM and MoToM were unrelated in this study, the previous work in this area is mixed; prior studies have found that ToM precedes MoToM (Loureiro & Souza, 2013), that MoToM predicts ToM (Knobe, 2003a), that the two capacities emerge simultaneously (Leslie, Knobe, & Cohen, 2006) or have bidirectional influences (Smetana et al., 2012). Furthermore, as previously discussed, the current study found the novel finding that three-year-olds may not be successful at this task, and so this likely influenced the relation between this construct and ToM. It is also possible that when controlling for age and verbal skills (as was done here), the relation between ToM and MoToM is no longer evident. Regardless, this relation is still unclear, and should be studied at length in future studies.

The study under examination provided a number of new insights to the field of developmental psychology, which will be discussed as follows: first, a review and interpretation of the path analysis results; then, a discussion of applied moral-cognitive development within
early childhood; lastly, refocusing the work from moral and social-cognition from aggression to prosociality.

**Review of Proposed Model and Interpretation**

Again, the proposed model was a good fit to the data of the strict MoToM sample constraint using teacher-reported data for behavioral outcomes, indicating that the assemblage of constructs within the theoretically and empirically supported outline may be an accurate perspective, but also when special consideration is given to the measurement of applied moral cognition. However, the individual path estimates were largely non-significant. While this can be considered problematic under traditional methods of estimation, that may not be the case in path analysis: “there is some sense in SEM that the view of the whole model has precedence over that of individual effects” (p. 16, Kline, 2005). That is, testing for individual path significance is not the primary goal of path analysis, and therefore non-significant path estimates should not hold greater weight than good model fit.

Of course, given the measurement issue with the MoToM resulting in a constrained sample, there could be an issue of lack of power due to small sample size (Bollen, 1989). This could be considered problematic as this increases the risk of committing a Type II error (obtaining a non-significant p-value when it should be significant), which – in terms of path analysis and the Chi-square fit index – could result in a non-significant Chi-square value, indicating good fit. In other words, when using path analysis and SEM in general, restricting the sample size can in itself result in an acceptable Chi-square value ($p>.05$). However, the other fit indices – RMSEA and CFI – also indicated that the model was a good fit to the data, and so it is less likely that the non-significant Chi-square was obtained in error. Therefore, although the sample was constrained and there was a loss of power, the results suggest that the proposed
model was a good fit to the data. By the model fitting the data, this indicates that the presumed relationships as outlined throughout this manuscript are not contradicted, and furthermore, may be valid (Bollen, 1989).

Therefore, the non-significant path estimates may indicate that while individual constructs might not be major contributing factors, a more holistic perspective might be more appropriate. That is, the non-significant path estimates in the face of a good fitting model may indicate that each relation incrementally contributes to the structure at large. In terms of development, the current findings indicate that as inhibitory control and Theory of Mind skills develop, and as children become capable of recognizing and processing contradictory information in regards to others’ actions and intentions, they seem to change their own behavior upon reflection of such knowledge. This finding is in line with previous studies on aggression by Boerger (2015) who found that for children high in ToM, moral identity influenced their aggressive behavior, but this was not true for children who were low in ToM; however, Boerger failed to find any changes in prosocial behavior, did not examine inhibitory control abilities, and examined these constructs in a sample of older children (8- to 14-years-old). The current study therefore expands on this previous finding by examining these developmental facets in early childhood, at time at which most of these abilities first develop, and also by including inhibitory control abilities, which previous research shows is important for consideration in behavioral changes during the preschool years (Baker, Tisak, Tisak, & Jensen, under review).

**Applied Moral Cognition within Early Childhood**

In regards to the MoToM scoring, it would seem that examining this phenomenon within an early childhood sample comes with some difficulty; most noticeably, the youngest children provided inconsistent responses and – when asked for justification – were often not able to
provide any justification whatsoever. This finding is in fact in agreement with past research (Kristen, Sodian, Licata, & Killen, 2015) and may speak to the developmental nature of this ability by way of several underlying assumptions – the first focusing on cognitive ability, and the second on moral development.

The first assumption underlying the developmental trajectory of morally relevant ToM is that children must become cognitively capable of dissociating intent from outcome, and responding thusly. This assumption is in line with research on executive functioning capacity throughout preschool (Sabbagh, 2006), and is also examined in studies on ToM development (Ding & Li, 2004; Hughes & Ensor, 2008). As with the Stroop task (Gerstadt, Hong, & Diamond, 1994), children must be able to ignore a prepotent response in order to attend to, and reflect on, less automatic information. For example, consider that the MoToM task presents children with conflicting information: a neutral or positive intent with a negative outcome. In circumstances such as these younger children still attend to the outcome more so than the intent, which mirrors early supposition by Piaget (1965) outlining the tendency for children in the concrete operational stage of cognitive development to focus on tangible, physical aspects of reality. This first assumption would highlight that young children are not cognitively capable of dissociating a harmful outcome from a non-negative intent, and thus are only capable of processing the negative outcome, at least initially. With time and brain maturation, children can at least begin to develop the ability to process non-physical information, such as intent, but also the dueling pieces of information.

Next, the second assumption focuses on moral development: understanding that another’s intention is a discrete phenomenon apart from judgment of wrongness (Killen et al. (2011); Cushman, Sheketoff, Wharton, & Carey, 2013). More simply, this assumption stipulates that
judgment of an actor can be separated from the judgment of the act. Killen and colleagues (2011) and Cushman and colleagues (2013) found that later in early childhood there seems to be a shift whereby children progressively rely more on actor intent rather than negative outcome when deciding punishment (intent-based, or “subjective” moral judgment, versus outcome-based or “objective”) or are finally able to exploit this ability (Margoni & Surian, 2016). Hallmark work by Piaget (1965) – in line with his ideas of cognitive development – also illustrates a version of this: when a character accidentally created a large ink stain, younger children viewed this as morally worse than a character who intentionally created a small stain, whereas older children made the opposite judgment. That is, younger children focused their moral judgment on the outcome, and older children focused more on the intent. This assumption implies that as children develop they associate their internal representations of “badness,” punishment, and wrongness, associations that are oftentimes taught to children unintentionally (Cushman et al., 2013), and therefore that this must be unlearned in order to understand intention.

In sum, the finding on MoToM in 3-year-olds expands the existing work on MoToM in early childhood. The few studies using this recently developed paradigm focus on slightly older samples (e.g., 4- to 7-year-olds, Fu, Xiao, Killen, & Lee, 2014), while the original study – although calling the youngest cohort “3.5 year olds” – included as their youngest group a sample of children with a mean age of 49 months (Killen et al., 2011). The current finding that 3-year-olds seemed to not be capable of responding consistently may imply that this ability might not yet be developed to the point where children can dually process conflicting information and, further, cannot explicitly express this understanding upon multiple requests. This could suggest that young children have the ability to understand a supposed innateness of these actions per a “gut feeling” (Rottman & Young, 2015) that they may not yet be capable of explaining. Similar
Findings can be found with adults using much more elaborate paradigms (e.g., the trolley vignette, Greene, 2013), which again may indicate an innate “moral faculty” (Dwyer, 1999).

**Findings for Prosociality**

Much previous work on applied moral cognition in early childhood focuses on negative issues, such as harm (e.g., Killen et al., 2011; Smetana et al., 2012); and no studies to date have focused on applied moral cognition in regards to developmental changes in behavioral outcomes, with the exception of one study which found relations for aggression but not prosociality (Boerger, 2015). To understand the findings of the current study, therefore, requires some conjecture.

A direct interpretation (in line with work on negative outcomes: Cushman et al., 2013; Margoni & Surian, 2016) would be that by understanding neutral intent with negative outcomes children may be able to generalize this reasoning such that there can be neutral intent with *any* outcome, including positive outcomes, or – equally likely – that intent and outcome can be dissociated entirely. This interpretation is in line with the *constructivist* process (Rottman & Young, 2015), whereby children become increasingly adept at processing morally relevant situations by developing a continuous evaluation of social situations; this theory is largely supported by social domain theory (Smetana, 2006; Turiel, 1983). According to these theories, then, children use social situations to enhance their reasoning skills and apply these skills differentially in the case of social situations based on the context of the situation (Turiel, 1983). Recall that domain theory hypothesizes that social interactions can fall into a number of domains: moral, conventional, and personal (Tisak et al., 2006); and that reasoning processes differ among these. This social learning mechanism by which children learn to generalize their
experiences, then, could be limited within a particular domain, or could extend to other domains depending upon context (Rottman & Young, 2015).

**Extension to Developmental Changes in Behavioral Outcomes**

The only study of which I am aware that examined behavioral outcomes in a similar context (Boerger, 2015) found that, for those high in ToM, moral identity predicted aggressive behavior. Specifically, Boerger (2015) found that children who are quite adept at mental state understanding and identify with more positive moral attitudes are less likely to be aggressive, whereas those with more negative moral attitudes show greater amounts of aggression, but again only for those with high mental state understanding. The current study likewise found moderation by applied moral cognition only for those high in mental state understanding, but for prosocial behaviors, not aggressive ones.

Nevertheless, combining the findings of the current study with those of Boerger (2015) it would seem that developed mental state understanding is insufficient in predicting behavioral changes during child development. Within the scope of social-cognitive development, moral understanding and the ability to infer others’ moral intent seem to influence the child’s selection of behavior, but only if the child is capable of understanding another’s mental state. To this end, if a child has underdeveloped social-cognitive abilities, the child’s moral reasoning may not influence their prosocial behavior, and the child may continue to be as prosocial as the context would recommend.

**Implications for Child Development**

Within the scope of explaining behavioral prosociality as a function of applied morality on mental state understanding, I suggest that children reflect on and alter their own behavior as a function of increasing social reasoning and social interaction. Specifically, as children become
better at understanding others’ intention (both positive and negative) they learn that others’ prosocial acts may not be associated with positive intentions. This may lead them to reflect upon their own actions more thoroughly – perhaps identifying times when their own action was incongruent with their intent – and adjust their behavior accordingly. Although Capputi, Lecce, Pagnin, and Banerjee (2012) found that increased ToM is related to greater prosocial behavior, this study was conducted with a middle childhood sample; therefore this sample likely did not include any children with low ToM. By focusing on older children, then, previous studies may have missed the intricacies between ToM development and moral cognition as it applies to social behavior.

Among younger children – as examined here – it may be that they view some aspects of prosocial behavior as convention (expectations of which can vary by culture, context, etc.), whereas aggressive behavior is arguably a moral transgression regardless of setting. This hypothesis could explain why we saw a decrease in prosocial behavior, rather than an increase in aggression; even young children identify that immoral behaviors (such as hitting) are wrong even when permitted by authority (Tisak, 1986) and when rules explicitly allow it (Smetana, Jambon, Conry-Murray, & Sturge-Apple, 2012). Furthermore, children are more accepting of transgressions against convention, which they view as arbitrary and somewhat flexible (Tisak et al., 2006). Therefore, children may be less prosocial if they view these expectations as social prescriptions, but not necessarily morally bound, and therefore fewer adherences to these rules could be acceptable.

Another alternative would be that while rules and perspectives on aggressive behaviors are prohibitive ("we don’t hit"), perspectives on prosocial behavior are often more in line with social approval, and young children oftentimes utilize hedonistic reasoning in social situations
(Eisenberg, Lennon, & Roth, 1983). It may be, therefore, that young children are less likely to commit aggressive acts so as to avoid punishment (which directly affects the child), but become less prosocial in such instances where being prosocial invokes a cost to them. For example, a child may be less likely to share his or her treat because this would result in him or her getting less, even though sharing would make the child’s mother happy. This explanation is equally likely as the first, but is more developmentally limited; while hedonistic reasoning is common during early childhood, with age children perceive social situations less egocentrically and are more likely to consider others’ perspectives (Eisenberg, Lennon, & Roth, 1983). For example, by elementary school most children reason through social situations regarding prosocial behavior based upon the needs of others even when this contrasts with their own needs (Eisenberg, Lennon, & Roth, 1983).

Limitations and Future Directions

The current study was not without flaw; first, as a function of data cleaning the sample was restricted, decreasing the statistical power of the analyses; and second, the goal of measuring sub-types of aggressive behavior was not accomplished with the current sample.

First, the study was lacking in statistical power due to removing cases for two reasons: (1) children not passing the control questions for the ToM battery, and (2) the younger children not understanding the MoToM task. Regarding removing children who did not pass the ToM control questions, although this decision is standard practice within the field and was outlined before data collection began this nevertheless resulted in the loss of approximately 50 cases. Furthermore, the majority of studies on ToM development – while mentioning the use of control questions – usually do not report the number of cases lost by this process, and so there was no way to anticipate the number of cases which would be lost. With regard to removing children
due to MoToM response inconsistencies, this outcome could not have been anticipated from recent studies. As this is a relatively new paradigm (Killen, et al., 2011) it has only been used a handful of times before this study and always with slightly older samples. This could imply that younger children (3-year-olds) do not possess morally relevant mental state understanding, which would agree with some previous research on MoToM development (Smetana et al., 2012), or it could be that this measure is too advanced for younger children and that simpler measures should be designed for this use. As an additional note, a number of studies using path analysis with young children often report smaller sample sizes and thusly suffer from a lack of power as well (e.g., Lewis, Sullivan, & Kim, 2015); although this does not address the weakness of the current study, it may speak to the difficulty of obtaining larger sample sizes within the field of early childhood. Nevertheless, these weaknesses can be addressed in future studies by obtaining a larger sample size and utilizing ongoing data cleaning.

The second issue with this study was the failed attempt to measure minor facets of aggressive behavior – specifically, proactive-physical, proactive-relational, reactive-physical, and reactive-relational aggression. Although prior attempts did yield valid and important findings toward this endeavor (Baker, Tisak, Tisak, & Jensen, under review), the current extension did not allow for this conceptualization. To that end, there are a number of differences between the previous study and the current project. Although recruitment techniques were identical between the two studies, the current study resulted in a sample of particular affluence compared to the norms for the geographical area and in comparison to the prior study. Furthermore, while reliability values of the child self-report scales of aggression were quite low, they are actually higher than some studies using child self-report with young children (Edelbrock, Costello, Dulcan, Kalas, & Conover, 1985) and so the values found here are somewhat impressive.
Nonetheless, the low reliability could account for the lack of findings when using the child self-report in analyses, and therefore these tasks should be developed further.

The project presented here would have benefited from the addition of other considerations, such as other measures of behavioral outcomes and implications of behavioral outcomes. Although we obtained child self-report and teacher-report data, observational records could have provided a more reliable measure of social behaviors. Furthermore, incorporating the children’s perceptions of peer group selection, specific play activity selection within the group, and justifications for these selections would have allowed examination of the next logical extension of this study. This exploration would be especially pertinent given that past research has found differences in judgments of moral transgressions based on the relationship between the child and the target (Dunn, Cutting, & Demetriou, 2000; Slomkowski & Killen, 1992).

**Conclusions**

The study presented here used theoretical framing by Bandura (1986) and Turiel (1983; 2006), which allowed for consideration of cognitive, social, and behavioral components of early child development. Although the study may have been underpowered, the fit indices suggest that the proposed model was a good fit to the data, whereby inhibitory control, ToM, moral judgment, and MoToM incrementally influence aggressive and prosocial behaviors. The primary finding of the present study was that for children with developed mental state understanding (ToM), applied moral cognition was negatively related to prosocial behavior (greater understanding of others’ motives predicted less prosocial behavior) but for children with low mental state understanding there existed no relation. This implies that once children are capable of understanding others’ individual beliefs, desires, expectations, etc., they begin to construct moral reasoning processes and then use these processes to regulate and alter their own behavior.
by way of self-reflection. This extends prior supposition on social and cognitive development by showing that applied moral reasoning affects behavior selection in preschool-age children as a function of mental state understanding. Future studies should examine how this change in behavior selection influences peer group selection and play time activities.


http://dx.doi.org/10.1037/h0100133


blockage: Its antecedent causes and its effect on toddler persistence. *Developmental 
Psychology, 51*, 1553-63.

10.1037/0012-1649.44.2.523

Loeber, R., & Hay, D. (1997). Key issues in the development of aggression and violence from 
childhood to early adulthood. *Annual Review of Psychology, 48*, 371-410. doi: 
10.1146/annurev.psych.48.1.371

Lonigro, A., Laghi, F., Baiocco, R., & Baumgartner, E. (2014). Mind reading skills and empathy: 
evidence for nice and nasty ToM behaviors in school-aged children. *Journal of Child 

432723542011311

determination of sample size for covariance structure modeling. *Psychological Methods, 
1*, 130-149.


Pratt, C., & Bryant, P. (1990). Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development, 61,* 973-982. doi: 10.1111/j.1467-8624.1990.tb02835.x


http://search.proquest.com/docview/304699226?accountid=26417


Table 1.

Descriptive statistics and demographics by age.

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>2 year old (N = 1)</th>
<th>3 year olds (N=56)</th>
<th>4 year olds (N=73)</th>
<th>5 year olds (N=39)</th>
<th>6 year olds (N=7)</th>
<th>Total (N=176)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Age in months (SD)</td>
<td>33</td>
<td>42.8 (0.39)</td>
<td>53.3 (0.43)</td>
<td>64.7 (0.59)</td>
<td>74.9 (0.60)</td>
<td></td>
</tr>
<tr>
<td>Parentage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living with mother only</td>
<td>-</td>
<td>16% (N=9)</td>
<td>8% (N=6)</td>
<td>10% (N=4)</td>
<td>29% (N=2)</td>
<td>11.9% (N=21)</td>
</tr>
<tr>
<td>Living with father only</td>
<td>-</td>
<td>-</td>
<td>1% (N=1)</td>
<td>3% (N=1)</td>
<td>-</td>
<td>1.1% (N=2)</td>
</tr>
<tr>
<td>Living with both parents</td>
<td>100% (N=1)</td>
<td>84% (N=47)</td>
<td>89% (N=65)</td>
<td>87% (N=34)</td>
<td>71% (N=5)</td>
<td>86.4% (N=152)</td>
</tr>
<tr>
<td>Mean Parents’ Education Level*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>4.53 (.18)</td>
<td>4.63 (.14)</td>
<td>4.82 (.23)</td>
<td>5.43 (.30)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father</td>
<td>4.30 (.18)</td>
<td>4.32 (.17)</td>
<td>4.47 (.25)</td>
<td>5.71 (.29)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender (% male/% female)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>-</td>
<td>64% (N=35)</td>
<td>48% (N=35)</td>
<td>56% (N=22)</td>
<td>57% (N=4)</td>
<td>54.5% (N=96)</td>
</tr>
<tr>
<td>Female</td>
<td>100% (N=1)</td>
<td>38% (N=21)</td>
<td>52% (N=38)</td>
<td>44% (N=17)</td>
<td>43% (N=3)</td>
<td>45.5% (N=80)</td>
</tr>
<tr>
<td>Mean Family Income</td>
<td>$110,530 ($8,390)</td>
<td>$123,275 ($12,457)</td>
<td>$118,918 ($21,663)</td>
<td>$101,328 ($17,867)</td>
<td>$116,093</td>
<td></td>
</tr>
</tbody>
</table>

Note: For parents’ education: 1 = some high school, 2 = high school diploma, 3 = some college, 4 = college graduate, 5 = some graduate school, 6 = graduate degree.
Table 2.

*Individual school sample size and demographics.*

<table>
<thead>
<tr>
<th>Location</th>
<th>Community Affiliation</th>
<th>N</th>
<th>Mean Income</th>
<th>Std. Dev.</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>School 1</td>
<td>Sub-rural</td>
<td>15</td>
<td>98,891</td>
<td>67,020</td>
<td>19,200-300,000</td>
</tr>
<tr>
<td>School 2</td>
<td>Sub-rural</td>
<td>16</td>
<td>61,812</td>
<td>29,678</td>
<td>21,000-110,000</td>
</tr>
<tr>
<td>School 3</td>
<td>Sub-rural</td>
<td>16</td>
<td>78,883</td>
<td>31,143</td>
<td>40,000-145,000</td>
</tr>
<tr>
<td>School 4</td>
<td>Suburban</td>
<td>12</td>
<td>82,412</td>
<td>32,129</td>
<td>33,000-145,000</td>
</tr>
<tr>
<td>School 5</td>
<td>Urban</td>
<td>3</td>
<td>37,855</td>
<td>13,346</td>
<td>23,000-50,000</td>
</tr>
<tr>
<td>School 6</td>
<td>Sub-rural</td>
<td>19</td>
<td>109,500</td>
<td>33,952</td>
<td>60,000-200,000</td>
</tr>
<tr>
<td>School 7</td>
<td>Sub-rural</td>
<td>4</td>
<td>94,750</td>
<td>31,255</td>
<td>70,000-140,000</td>
</tr>
<tr>
<td>School 8</td>
<td>Sub-rural</td>
<td>9</td>
<td>100,028</td>
<td>25,686</td>
<td>77,000-160,000</td>
</tr>
<tr>
<td>School 9</td>
<td>Suburban</td>
<td>12</td>
<td>142,405</td>
<td>54,558</td>
<td>65,000-250,000</td>
</tr>
<tr>
<td>School 10</td>
<td>Urban</td>
<td>4</td>
<td>81,728</td>
<td>5,765</td>
<td>73,801-86,000</td>
</tr>
<tr>
<td>School 11</td>
<td>Suburban</td>
<td>9</td>
<td>121,975</td>
<td>34,883</td>
<td>80,000-200,00</td>
</tr>
<tr>
<td>School 12</td>
<td>Suburban</td>
<td>27</td>
<td>159,047</td>
<td>157,047</td>
<td>315,000-900,000</td>
</tr>
<tr>
<td>School 13</td>
<td>Suburban</td>
<td>30</td>
<td>160,085</td>
<td>150,267</td>
<td>46,000-750,000</td>
</tr>
</tbody>
</table>

**Total Sample** | **176** | **116,093** | **98,628** | **19,200-900,000**
Table 3.

*Theory of Mind battery.*

<table>
<thead>
<tr>
<th>Task</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diverse Desires</td>
<td>Child judges that two persons (child vs. someone else) have different desires about the same objects.</td>
</tr>
<tr>
<td>Knowledge Access</td>
<td>Child sees what is in a box and judges the knowledge of another person who does not see what is in the box.</td>
</tr>
<tr>
<td>Appearance-Reality</td>
<td>Child judges what another person will think an object actually is given that person’s false belief about its appearance.</td>
</tr>
<tr>
<td>Explicit False Belief</td>
<td>Child judges how a person will search for an object, given that person’s mistaken belief of the location of the object.</td>
</tr>
<tr>
<td>Real-Apparent Emotion</td>
<td>Child judges that a person can feel one thing but display a different emotion.</td>
</tr>
</tbody>
</table>
Table 4.

*Constructs and measures.*

<table>
<thead>
<tr>
<th>Construct</th>
<th>Measurement</th>
<th>Reporter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theory of Mind</td>
<td>Diverse Desires</td>
<td>Child</td>
</tr>
<tr>
<td></td>
<td>Knowledge Access</td>
<td>Child</td>
</tr>
<tr>
<td></td>
<td>Appearance Reality</td>
<td>Child</td>
</tr>
<tr>
<td></td>
<td>Explicit False Belief</td>
<td>Child</td>
</tr>
<tr>
<td></td>
<td>Real-Apparent Emotion</td>
<td>Child</td>
</tr>
<tr>
<td>Moral Theory of Mind</td>
<td>MoToM</td>
<td>Child</td>
</tr>
<tr>
<td>Morality</td>
<td>Moral Transgressions</td>
<td>Child</td>
</tr>
<tr>
<td>Inhibitory Control</td>
<td>Stroop – Response Time</td>
<td>Child</td>
</tr>
<tr>
<td></td>
<td>Stroop – Accuracy</td>
<td>Child</td>
</tr>
<tr>
<td>Aggression</td>
<td>Aggressive Behaviors Interview</td>
<td>Child</td>
</tr>
<tr>
<td></td>
<td>Aggressive Behaviors Questionnaire – Teacher Form</td>
<td>Teacher</td>
</tr>
<tr>
<td>Prosocial</td>
<td>Prosocial Behaviors Interview</td>
<td>Child</td>
</tr>
<tr>
<td></td>
<td>Prosocial Behaviors Questionnaire – Teacher Form</td>
<td>Teacher</td>
</tr>
</tbody>
</table>
Table 5.

Demographic statistics for the five sample compositions.

<table>
<thead>
<tr>
<th></th>
<th>Full Sample (N=176)</th>
<th>ToM Control (N=121)</th>
<th>MoToM Restriction 1 (N=109)</th>
<th>MoToM Restriction 2 (N=84)</th>
<th>Age Restricted (N=110)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (months)</strong></td>
<td>53.24 (9.83)</td>
<td>54.34 (9.68)</td>
<td>54.80 (9.67)</td>
<td>56.23 (9.26)</td>
<td>55.85 (8.79)</td>
</tr>
<tr>
<td>Range</td>
<td>[33-78]</td>
<td>[37-78]</td>
<td>[37-78]</td>
<td>[40-78]</td>
<td>[42-78]</td>
</tr>
<tr>
<td><strong>Gender (% male)</strong></td>
<td>54.5%</td>
<td>51.2%</td>
<td>51.4%</td>
<td>52.4%</td>
<td>51.8%</td>
</tr>
<tr>
<td><strong>Mother’s Ed. mean</strong></td>
<td>4.67</td>
<td>4.59</td>
<td>4.60</td>
<td>4.67</td>
<td>4.61</td>
</tr>
<tr>
<td><strong>Father’s Ed. mean</strong></td>
<td>4.40</td>
<td>4.29</td>
<td>4.34</td>
<td>4.27</td>
<td>4.34</td>
</tr>
<tr>
<td><strong>Income (mean)</strong></td>
<td>116,093</td>
<td>112,864</td>
<td>114,673</td>
<td>118,104</td>
<td>115,760</td>
</tr>
</tbody>
</table>

*Note:* For parents’ education: 1 = some high school, 2 = high school diploma, 3 = some college, 4 = college graduate, 5 = some graduate school, 6 = graduate degree.
Table 6.

*Item measure summary and fit statistics for 5-item Rasch Model.*

<table>
<thead>
<tr>
<th>Task</th>
<th>Item Location</th>
<th>$X^2$</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Diverse Desires</td>
<td>-1.66</td>
<td>8.66</td>
<td>4</td>
</tr>
<tr>
<td>2 Knowledge Access</td>
<td>-0.94</td>
<td>8.63</td>
<td>4</td>
</tr>
<tr>
<td>3 Explicit False Belief</td>
<td>-0.34</td>
<td>4.14</td>
<td>4</td>
</tr>
<tr>
<td>4 Real-Apparent Emotion</td>
<td>0.76</td>
<td>3.29</td>
<td>3</td>
</tr>
<tr>
<td>5 Supplemental False Belief</td>
<td>0.80</td>
<td>7.24</td>
<td>3</td>
</tr>
</tbody>
</table>

Note: all $p$ values > .05.
Table 7.


<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Teacher Report</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Physical Aggression</td>
<td>(.93)</td>
<td>.77**</td>
<td>-.64**</td>
<td>.09</td>
<td>.20*</td>
<td>-.19*</td>
<td>0.60</td>
<td>0.74</td>
</tr>
<tr>
<td>2. Relational Aggression</td>
<td>(.91)</td>
<td>-.49**</td>
<td>.07</td>
<td>.06</td>
<td>-.06</td>
<td>0.64</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>3. Prosocial Behavior</td>
<td>(.92)</td>
<td>-.10</td>
<td>-.14</td>
<td>.13</td>
<td>2.97</td>
<td>0.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Child Report</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Physical Aggression</td>
<td>(.60)</td>
<td>.32**</td>
<td>-.09</td>
<td>.33</td>
<td>-.20*</td>
<td>0.36</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>5. Relational Aggression</td>
<td>(.33)</td>
<td>-.20*</td>
<td>.77</td>
<td>1.95</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: Values on diagonal are reliability coefficients. Correlation values are based on standardized values. Means and SD are based on unstandardized.

*p<.05, **p<.01
Table 8.

Effects of age on IC-accuracy, IC-reaction time, ToM, MoToM, and behavioral outcomes.

<table>
<thead>
<tr>
<th>Outcome Variable</th>
<th>$R^2$</th>
<th>$B$</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. IC – Accuracy</td>
<td>.03</td>
<td>.08</td>
<td>3.72$^a$</td>
</tr>
<tr>
<td>2. IC – Reaction Time</td>
<td>.05</td>
<td>-.02</td>
<td>6.04**</td>
</tr>
<tr>
<td>3. ToM</td>
<td>.27</td>
<td>.06</td>
<td>44.87**</td>
</tr>
<tr>
<td>4. MoToM</td>
<td>.01</td>
<td>.01</td>
<td>0.79</td>
</tr>
<tr>
<td>6. Physical Aggression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher Report</td>
<td>.02</td>
<td>-.01</td>
<td>1.87</td>
</tr>
<tr>
<td>Child Self-report</td>
<td>.02</td>
<td>-.01</td>
<td>2.34</td>
</tr>
<tr>
<td>7. Relational Aggression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher Report</td>
<td>.00</td>
<td>.00</td>
<td>.52</td>
</tr>
<tr>
<td>Child Self-Report</td>
<td>.02</td>
<td>-.01</td>
<td>2.84$^a$</td>
</tr>
<tr>
<td>8. Prosocial</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher Report</td>
<td>.05</td>
<td>.02</td>
<td>5.65*</td>
</tr>
<tr>
<td>Child Self-Report</td>
<td>.17</td>
<td>.03</td>
<td>24.63**</td>
</tr>
</tbody>
</table>

$^a p<.10$  $^* p<.05$  $^{**} p<.01$

Note: for analyses using Teacher Report, df = (1,116); for all others: df = (1,119).
Table 9.

Means for all variables of interest by gender within the conservative MoToM restricted sample.

<table>
<thead>
<tr>
<th></th>
<th>Males (N=44)</th>
<th>Females (N=40)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IC-Response Time</td>
<td>1.88 (0.74)</td>
<td>2.13 (0.70)</td>
</tr>
<tr>
<td>IC-Accuracy</td>
<td>9.93 (4.57)</td>
<td>9.97 (5.07)</td>
</tr>
<tr>
<td>ToM</td>
<td>2.68 (1.13)</td>
<td>2.52 (1.06)</td>
</tr>
<tr>
<td>MJ</td>
<td>1.32 (0.74)</td>
<td>1.48 (0.78)</td>
</tr>
<tr>
<td>MoToM</td>
<td>1.50 (1.02)</td>
<td>1.72 (1.11)</td>
</tr>
<tr>
<td><strong>Aggression</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher Report</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical</td>
<td>0.53 (0.69)</td>
<td>0.51 (0.70)</td>
</tr>
<tr>
<td>Relational</td>
<td>0.54 (0.57)</td>
<td>0.68 (0.71)</td>
</tr>
<tr>
<td>Child Report</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical</td>
<td>0.30 (0.68)</td>
<td>0.26 (0.61)</td>
</tr>
<tr>
<td>Relational</td>
<td>0.31 (0.56)</td>
<td>0.26 (0.45)</td>
</tr>
<tr>
<td><strong>Prosocial</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher Report</td>
<td>3.05 (0.57)</td>
<td>3.07 (0.66)</td>
</tr>
<tr>
<td>Child Report</td>
<td>2.05 (0.76)</td>
<td>2.25 (0.71)</td>
</tr>
</tbody>
</table>

Note: Response time was measured in seconds. Accuracy range of possible scores [0-16]. ToM range of possible scores [0-5]. MJ range of possible scores [0-4]. MoToM range of possible scores [0-4]. Teacher reports (aggression and prosocial) range of possible scores [0-4]. Child self-report (aggression and prosocial) range of possible scores [0-3].
Figure 1.

*Hypothesized relations between primary constructs.*
Figure 2.

Graded scale used in MoToM and Moral Transgression tasks.

(Killen, Mulvey, Richardson, Jampol, & Woodward, 2011)
Figures 3a and 3b.

*Night/Day Stroop Task Card Images.*

(Gerstadt, Hong, Diamond, 1994).
Figure 4.

Path Model: Teacher Data.

Restricted Sample: $X^2(24, N=84) = 34.24, p = .08$; CFI = 0.97; RMSEA = 0.07 [0.00, 0.12]

Note: Age was controlled for in all variables, and verbal scores were controlled on ToM and the ToM-by-MoToM interaction term. Error terms were included for all endogenous variables, and were allowed to correlate for ToM, MoToM, and ToM-by-MoToM, as well as for the three outcome variables. Control variables and error terms were not included in the figure for simplicity. Exogenous variables were allowed to correlate.
Figure 5.

Path Model: Child Data.

Restricted Sample: $X^2(26, \text{N}=84) = 49.48$, $p < .05$; CFI=0.91; RMSEA = 0.11; 90% CI [0.07, 0.16].

Note: Age was controlled for in all variables, and verbal scores were controlled on ToM and the ToM-by-MoToM interaction term. Error terms were included for all endogenous variables, and were allowed to correlate for ToM, MoToM, and ToM-by-MoToM, as well as for the three outcome variables. Control variables and error terms were not included in the figure for simplicity. Exogenous variables were allowed to correlate.
Moderation of MoToM on the relation from ToM on teacher-reported prosocial behavior, whereby greater mental state understanding was predictive of less prosocial behavior, but only for children who indicated understanding the target character’s motive in the MoToM story.
APPENDIX A. AGGRESSIVE BEHAVIORS QUESTIONNAIRE – TEACHER FORM.

Please think about each question for a moment, and then indicate **how often** the child in question exhibits each behavior.

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>never</td>
<td>seldom</td>
<td>sometimes</td>
<td>often</td>
<td>always</td>
</tr>
</tbody>
</table>

1. When others pinch or shove this child, he will destroy their things, like artwork or a favorite toy.
2. Hits or pushes other children when unprovoked.
3. Will say something mean to a child to hurt that child’s feelings.
4. When this child’s feelings are hurt, this child will be mean to that classmate.
5. In order to be mean, he will verbally threaten another child with physical harm.
6. This child will tear up another’s artwork for fun.
7. If a classmate is mean to this child, he or she responds will refuse to play with that child later.
8. When this child doesn’t get what he wants, he will threaten another child.
9. If pushed by other children, this child will respond by hitting or pushing.
10. This child lies about another child’s actions just to be mean.
11. Tries to get friends and classmates to dislike a particular student just for fun.
12. Unprovoked, this child pulls hair or shoves.
13. When feeling rejected by another child, this child will make other students dislike that child.
14. Tells lies about another child in response to being excluded from a game.
15. When this child is hit, this child will hit back.
16. In order to make a particular child feel bad, this child might refuse to play with that child.

*Note 1: Physical-Proactive items: #2, 5, 6, 12
   Physical-Reactive items: #1, 8, 9, 15
   Relational-Proactive items: #3, 10, 11, 16
   Relational-Reactive items: #4, 7, 13, 14*
### APPENDIX B. AGGRESSIVE BEHAVIORS – CHILD INTERVIEW

<table>
<thead>
<tr>
<th>Type of Aggression</th>
<th>Original item</th>
<th>Adapted Item for child (Opposing Puppet Statement)</th>
</tr>
</thead>
</table>
| #1 Physical-Reactive | When others pinch or shove this child, he will destroy their things, like artwork or a favorite toy. | Prompt: What do you do when someone hurts you?  
P1: I rip up their stuff!  
P2: I don’t rip up their stuff! |
| #2 Physical-Proactive | Hits or pushes other children when unprovoked. | It’s funny to hit other kids!/It’s not fun to hit other kids! |
| #6 Physical-Proactive | This child will tear up another’s artwork for fun. | I like to rip up other kids’ stuff!/I don’t like to rip up other kids stuff! |
| #8 Physical-Reactive | When this child doesn’t get what he wants, he will threaten another child. | Prompt: What do you do when you’re mad?  
P1: I say I’m gonna hit somebody!  
P2: I don’t say I’m gonna hit somebody! |
| #11 Relational-Proactive | Tries to get friends and classmates to dislike a particular student just for fun. | I like when me and my friends pick on other kids!/I don’t like when me and my friends pick on other kids! |
| #13 Relational-Reactive | When feeling rejected by another child, this child will make other students dislike that child. | Prompt: When someone won’t play with you, what do you do?  
P1: I tell my friends to be mean to him!  
P2: I don’t tell my friends to be mean! |
| #14 Relational-Reactive | Tells lies about another child in response to be excluded from a game. | Prompt: When someone won’t let you play, what do you do?  
P1: I get them in trouble!  
P2: I don’t get them in trouble! |
| #16 Relational-Proactive | In order to make a particular child feel bad, this child might refuse to play with that child. | I think it’s fun when I don’t let other kids play with me/I think it’s fun to let all the kids play. |
APPENDIX C. PROSOCIAL BEHAVIORS QUESTIONNAIRE – TEACHER FORM.

(Ladd & Profilet, 1996)

Please think about each question for a moment, and then indicate **how often** the child in question exhibits each behavior.

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>never</td>
<td>seldom</td>
<td>sometimes</td>
<td>often</td>
<td>always</td>
<td></td>
</tr>
</tbody>
</table>

1. Seems concerned when classmates are distressed.
2. Is kind toward classmates.
3. Listens to classmates.
4. Compromises in conflicts with classmates.
5. Is cooperative with classmates.
7. Shows concern for moral issues (e.g., fairness, welfare of others).
8. Offers help or comfort when classmates are upset.
**APPENDIX D. PROSOCIAL BEHAVIORS – CHILD INTERVIEW.**

<table>
<thead>
<tr>
<th>Original Item</th>
<th>Adapted for Child (Opposing Puppet Statement)</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1 Seems concerned when classmates are distressed.</td>
<td>When someone is sad, I’m sad too./When someone is sad, I don’t get sad.</td>
</tr>
<tr>
<td>#2 Is kind toward classmates</td>
<td>I like being nice/I don’t like being nice</td>
</tr>
<tr>
<td>#3 Listens to classmates</td>
<td>I like to listen to other kids/I don’t like to listen to other kids</td>
</tr>
<tr>
<td>#4 Compromises in conflicts with classmates</td>
<td>I like when everybody gets a turn/I don’t like to take turns</td>
</tr>
<tr>
<td>#5 Is cooperative with classmates</td>
<td>I like working with the other kids/I don’t like working with the other kids</td>
</tr>
<tr>
<td>#6 Is friendly toward classmates</td>
<td>I like playing with my friends/I don’t like playing with my friends</td>
</tr>
<tr>
<td>#7 Shows concern for moral issues</td>
<td>I don’t like when other kids are mean/I don’t care if other kids are mean</td>
</tr>
<tr>
<td>#8 Offers help or comfort when classmates are upset</td>
<td>I like helping when someone is sad/I don’t like to help if someone is sad</td>
</tr>
</tbody>
</table>
APPENDIX E. PARENT CONSENT FORM.

Dear Parent or Guardian:

Introduction: I am a graduate student of psychology at Bowling Green State University working with Dr. Marie S. Tisak, Professor of Psychology at Bowling Green State University. We’re interested in how preschool children understand everyday social interactions.

Purpose & Benefits: In this study we are interested in preschool-aged children’s thinking about what other people know, how children view “right and wrong”, and their social behaviors in the classroom. Past research has only focused on kids’ thinking of others’ knowledge with certain types of classroom behavior. This study looks at children’s different classroom behaviors, such as sharing, helping, and hitting. This information will help us understand young children’s ideas about what other people think and how that might influence their behavior. Also, we have found that preschoolers enjoy participating in these games, and have fun playing them. All children also receive a small prize at the end as a reward.

Procedure: We will be showing the children two puppet shows, and asking them what they think each puppet was thinking, and which puppet they were more like. We will also be playing a card game, as well as showing the children pictures of everyday objects (e.g., broccoli, a swing set), and asking the children how people think and feel about the objects. Vocabulary skills will also be measured. A researcher will interview each child individually two times, and each interview will not last longer than 15 minutes. We will also be asking the teacher to fill out a questionnaire about the child’s typical classroom behavior.

Voluntary nature: We are asking your permission for your son or daughter to participate in the study. Participation is completely voluntary. Please be assured that regardless of your decision, this will not have any impact on your relationship with the preschool or Bowling Green State University. Following your permission the study will be described to your child. If your child wishes to participate, he or she will then be interviewed. You may withdraw your consent at any time and your child can choose to stop participating at any time for any reason. Also, regardless of whether or not your child participates, this will have no impact on his/her preschool or with Bowling Green State University.

Risks: The risks for your child to participate are no greater than those encountered in daily life.

Confidentiality: Your child’s answers and comments will be kept confidential. My interest is not in a particular child’s thinking but in an understanding of children as a group. To ensure confidentiality, all answer sheets will be marked with an ID number only. This ID number will be used in analyzing the information, rather than the individual’s name or other identifiers. Any identifying information will be destroyed at the completion of the study. All documents will be locked in a file cabinet in my advisor’s office, where only my advisor and I will have access. When the project is completed, a report will be sent to your child’s school where it will be made available for you upon request.

Contact Information: If you have any questions concerning any part of this research, please feel free to contact me at (419) 372-4395 or my advisor at (419) 372-2273, or by e-mail: ebaker@bgsu.edu, or mtisak@bgsu.edu. You may also contact the Human Subjects Review Board at (419) 372-7716 or by e-mail: hsrb@bgsu.edu, if you have any questions about your child’s rights as a participant in this research.

What you need to do: Please indicate on the next page your decision regarding your child’s participation. If you choose to let your child participate, please complete the demographic information. Place the completed form in the envelope provided, and give the envelope to your child so that he/she can return them to school. Thank you very much for your time and consideration.

Sincerely,
Erin R. Baker, M.S.
Doctoral Candidate

Marie S. Tisak, Ph.D.
Professor of Psychology

206 Psychology Building
Bowling Green, OH 43403-0232

419-372-2301
Fax: 419-372-6513

BGSU IRB 
APPROVED FOR USE
IRBNet ID #: #015316
EFFECTIVE: 08/12/2015
EXPIRES: 08/12/2018
Parental Consent Form
Study on Preschool Children’s Thinking and Classroom Behaviors

☐ By checking this box, I am indicating that I have read the attached letter.

At this time: ________ I give permission for my child to participate.

________ I do not give permission for my child to participate.

If you do not give your child permission to participate, please fill out only your child’s name, and
give to your child to return to school. Thank you!

If you do give permission for your child to participate, please fill out the following information
below and place it in the envelope provided. Please return this to the school. Thank you!

Child’s Name (Please print) ______________________
Child’s School _____________________________
Hometown: _________________________________
Print name of parent or guardian ________________________
Signature of parent or guardian __________________________
Your relationship to this child __________________________
Today’s Date _________________________________

Basic background information:
1. Child’s Birthdate ____________ ☐ Child’s age ____________
2. What adults live with the child at home? (please circle all that apply)
   Mother  Father  Grandparent(s)  Adult Sibling(s)  Aunt(s)/Uncle(s)

3. How many other siblings does the child have (please circle), and what are their ages?

   None  1  2  3  4  5

   Ages: ______________________

4. Mother’s highest level of education (please circle):
   Some high school  High school diploma  Some college  College degree  Some graduate school  Graduate degree

5. Father’s highest level of education (please circle):
   Some high school  High school diploma  Some college  College degree  Some graduate school  Graduate degree

6. Approximate household income (before taxes): ______________________

7. What is the ethnic/racial background of this child? (Please Circle)
   Caucasian  African-American  Hispanic Asian-American  Multi-Racial  Other

   *If “Other” please specify: ______________________
APPENDIX F. CHILD ASSENT FORM.

Participant’s Verbal Assent Form

_______ Yes, the child would like to help with the project.

_______ No, the child would not like to help with the project.

Print Child’s Name

________________________

Initials of Researcher

________________________
APPENDIX G. HSRB APPROVAL FORM.

<table>
<thead>
<tr>
<th>DATE:</th>
<th>May 12, 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>TO:</td>
<td>Erin Baker</td>
</tr>
<tr>
<td>FROM:</td>
<td>Bowling Green State University Human Subjects Review Board</td>
</tr>
<tr>
<td>PROJECT TITLE:</td>
<td>[751367-2] Theory of Mind and Morality Development: Impacts on Aggression and Prosociality</td>
</tr>
<tr>
<td>SUBMISSION TYPE:</td>
<td>Revision</td>
</tr>
<tr>
<td>ACTION:</td>
<td>APPROVED</td>
</tr>
<tr>
<td>APPROVAL DATE:</td>
<td>May 12, 2015</td>
</tr>
<tr>
<td>EXPIRATION DATE:</td>
<td>May 5, 2016</td>
</tr>
<tr>
<td>REVIEW TYPE:</td>
<td>Expedited Review</td>
</tr>
<tr>
<td>REVIEW CATEGORY:</td>
<td>Expedited review category # 7</td>
</tr>
</tbody>
</table>

Thank you for your submission of Revision materials for this project. The Bowling Green State University Human Subjects Review Board has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a project design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

The final approved version of the consent document(s) is available as a published Board Document in the Review Details page. You must use the approved version of the consent document when obtaining consent from participants. Informed consent must continue throughout the project via a dialogue between the researcher and research participant. Federal regulations require that each participant receives a copy of the consent document.

Please note that you are responsible to conduct the study as approved by the HSRB. If you seek to make any changes in your project activities or procedures, those modifications must be approved by this committee prior to initiation. Please use the modification request form for this procedure.

You have been approved to enroll 300 participants. If you wish to enroll additional participants you must seek approval from the HSRB.

All UNANTICIPATED PROBLEMS involving risks to subjects or others and SERIOUS and UNEXPECTED adverse events must be reported promptly to this office. All NON-COMPLIANCE issues or COMPLAINTS regarding this project must also be reported promptly to this office.

This approval expires on May 5, 2016. You will receive a continuing review notice before your project expires. If you wish to continue your work after the expiration date, your documentation for continuing review must be received with sufficient time for review and continued approval before the expiration date.

Good luck with your work. If you have any questions, please contact the Office of Research Compliance at 419-372-7716 or hsrb@bgsu.edu. Please include your project title and reference number in all correspondence regarding this project.