TEN-MONTH-OLDS’ EVALUATIONS OF ACCIDENTAL AND INTENTIONAL ACTIONS

by

T. Doan Le

B.Sc., The University of Waterloo, 2011

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Psychology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2013

© T. Doan Le, 2013
Abstract

Mature moral judgments rely on the analysis of both the outcomes of others’ actions and the mental states that drive them. Past research has shown that when there is conflict between outcome and intention, young children rely on outcome information to evaluate others, while older children and adults privilege intention (Piaget, 1932/1965). This suggests that there is a shift from outcome-based to intention-based judgments occurring in development.

However, the current study suggests that even 10-month-old infants evaluate moral agents on the basis of their underlying mental states. Infants were presented with puppet shows in which a protagonist was either intentionally or accidentally helped or hindered. Infants were then given a forced choice between the accidental and intentional puppets. Results indicate that infants’ preference for the accidental versus the intentional character differed by condition $[\chi^2(1, N = 60) = 11.28, p < .001, \phi = .43]$; infants preferred intentional to accidental helpers (Binomial, $p < .05$), but preferred accidental to intentional hinderers (Binomial, $p < .05$). These results suggest that the capacity to evaluate others on the basis of intention arises much earlier on in development than previously suggested and contradicts earlier claims of a developmental shift from outcome- to intention-based judgments.
Preface

All research reported was conducted at UBC’s Centre for Infant Cognition, and was supervised by Dr. J. Kiley Hamlin. Ethics approval for this research was obtained from UBC’s Behavioral Research Ethics Board under the approval number H10-01808.
Table of Contents

Abstract ................................................................................................................................. ii

Preface ................................................................................................................................. iii

Table of Contents ................................................................................................................... iv

List of Tables .......................................................................................................................... vi

List of Figures ......................................................................................................................... vii

Acknowledgements .............................................................................................................. viii

1 Introduction ......................................................................................................................... 1

1.1 How do children and adults evaluate moral agents? ....................................................... 2

1.2 Infants’ understanding of mental states ........................................................................ 7

1.3 Infants’ third-party evaluations of morally-relevant actions ......................................... 9

1.4 Failed attempts vs. accidents .......................................................................................... 13

1.5 The present investigation ............................................................................................... 16

2 Experiment ......................................................................................................................... 18

2.1 Participants ....................................................................................................................... 18

2.2 Procedure ......................................................................................................................... 18

2.2.1 Familiarization ................................................................................................................. 19

2.2.1.1 Intentional helping .................................................................................................. 21

2.2.1.2 Accidental helping .................................................................................................. 22

2.2.1.3 Intentional hindering ............................................................................................... 23

2.2.1.4 Accidental hindering ............................................................................................... 24

2.2.2 Choice .......................................................................................................................... 25

2.2.3 Counterbalancing and reliability ................................................................................ 26
List of Tables

Table 1. Infants’ looking time towards events (s) separated by condition .....................28
List of Figures

Figure 1. Failed helping and failed hindering events (Hamlin, in press) ..................... 13
Figure 2. Critical actions displayed in puppet show events........................................... 27
Figure 3. Results of infants’ choice ............................................................................. 30
Acknowledgements

First and foremost, I would like to thank my wonderful supervisor, Kiley Hamlin, for her guidance and support. This has been an incredible experience and I am so grateful to have had this opportunity. Many thanks to Susan Birch and Mark Schaller for being on my thesis committee and for offering new ways to think about this research.

I would also like to extend thanks to Ori Friedman, who has been an amazing mentor over the years. His creativity and passion for research has been truly inspirational.

This research would not have been possible without the infants and families who participated, and the hardworking members of the Centre for Infant Cognition. In particular, thanks to Janine Gellerman for making sure that the lab ran smoothly and for making the lab a fun environment to work in; Meghan MacPherson, Olesha Ratther, Nadiu Hui, and Carolyn Baer for their lively spirits and eagerness to help in any way possible.

Finally, I would like to thank all of my friends and family for supporting me along the way—I would not have been able to do it without them.
1 Introduction

Moral judgments play a critical role in helping us navigate our social world. Our assessments as to whether or not individuals are morally responsible for producing particular outcomes impacts our social evaluations of them. Those who intentionally cause good outcomes (e.g., donating to charity, saving a drowning child) are evaluated positively, while those who intentionally cause bad outcomes (e.g., robbing a bank, committing murder) are evaluated negatively. The ability to make these assessments of others is critical in helping us determine who it is beneficial to align with and who it is important to avoid.

As adults, we judge others quickly and automatically (e.g., Allbright, Kenny, & Malloy, 1988; Ambady & Rosenthal, 1993; Todorov, Mandisodza, Goren, & Hall, 2005) and our initial judgments about others influence how we interpret their behaviors and evaluate them in future interactions (e.g., Darley & Fazio, 1980; Heyman & Gelman, 1999; Higgins, Rholes, & Jones, 1977; Imamoğlu, 1976; Kelley, 1950; Rabin & Schrag, 1999). For example, after positively evaluating others, we will be more inclined to continue evaluating them positively and will even interpret any of their future ambiguous actions as benevolent or prosocial. Similarly, if we initially hold negative evaluations of others, we will be more inclined to continue evaluating them negatively and will interpret their future ambiguous actions as hostile or antisocial. But how exactly do adults come to judge others so quickly and automatically? One possibility is that the ability to evaluate others is acquired over time and becomes automatized as we gain experience making these types of judgments (much like riding a bike). Another
possibility is that the capacity to evaluate others is a core ability that is built in, as it confers a survival-related advantage (i.e., we need to be able to distinguish friend from foe). However, a third and more plausible possibility is that both nature and nurture play a role in developing our socio-evaluative capabilities. Thus, it is important to discern how much of the ability is built in and how much experience is required to shape it. One way to get at this distinction is to study infants, as they have very minimal experience in the world. By studying infants, we can come closer to pinpointing when socio-evaluative abilities begin to emerge and we can determine the relative role of experience on the development of these abilities.

In the present investigation, I will demonstrate that 10-month-old infants display early nuanced socio-moral evaluations. This work will suggest that infants evaluate moral agents on the basis of their underlying mental states, and that infants use these evaluations to guide their social preferences in future interactions. But before turning to the study of infants, it is first critical to discuss how children and adults evaluate moral agents, so that we have a basic understanding of what “mature” evaluations look like.

1.1 How do children and adults evaluate moral agents?

Mature moral judgments rely not only on the analysis of the outcomes of an actor’s actions, but even more critically on the underlying mental states driving these actions (see Killen & Smetana, 2008 for a review; see also Cushman, 2008; Cushman, Sheketoff, Wharton, & Carey, 2013; Hamlin, in press; Leslie, Knobe, & Cohen, 2006; Malle, 1999; Nobes, Panagiotaki, & Pawson, 2009; Young, Cushman, Hauser, & Saxe,
Those who act intentionally should be held morally responsible for their actions and should be more deserving of praise, blame, reward, or punishment than those who act unintentionally (and are not morally responsible for their actions). To illustrate the significance of mental states on moral judgments, imagine the following two scenarios (taken from Young et al., 2007):

1. Grace and her friend are taking a tour of a chemical plant. When Grace goes over to the coffee machine to pour some coffee, Grace’s friend asks for some sugar in her coffee. There is white powder in a container by the coffee. The white powder is a poison left behind by a scientist. The container is labeled “toxic”, so Grace believes that the white powder is a poison. Grace puts the substance in her friend’s coffee. Her friend drinks the coffee and gets sick.

2. Grace and her friend are taking a tour of a chemical plant. When Grace goes over to the coffee machine to pour some coffee, Grace’s friend asks for some sugar in her coffee. There is white powder in a container by the coffee. The white powder is a poison left behind by a scientist. The container is labeled “sugar”, so Grace believes that the white powder is regular sugar. Grace puts the substance in her friend’s coffee. Her friend drinks the coffee and gets sick.

In both of these examples, Grace puts poison in her friend’s coffee, resulting in her friend getting sick. However, Grace can only be held morally responsible for her actions in the first case, where she believed the white powder was poison and thus, intentionally poisoned her friend. In the second scenario, Grace believed the powder was sugar, and thus only accidentally poisoned her friend. Consistent with these intuitions, adults rate the intentional harming as more blameworthy than the unintentional harming (Young et al., 2007). Importantly, these examples demonstrate
that the exact same actions can be viewed differently depending on an actor’s underlying mental states. Thus, an analysis of an actor’s mental states is critical in making mature moral judgments.

When there is a match between actors’ intentions and the outcomes that they produce, children’s and adults’ evaluations of others are clear cut—those who have bad intentions and produce bad outcomes are negatively evaluated, while those who have good intentions and produce good outcomes are positively evaluated. However, when outcome and intention conflict or when intention is ambiguous, young children and adults differ in their evaluations; young children rely more heavily on outcome information to evaluate others, while older children and adults privilege intention (e.g., Buchanan & Thompson, 1973; Gutkin, 1972; Imamoğlu, 1975; King, 1971; Piaget, 1932/1965; Yuill, 1984). For example, in Jean Piaget’s (1932/1965) moral reasoning task, children were told stories about two different transgressors: one who accidentally broke 15 cups, and one who intentionally broke one cup. When asked which of the two transgressors was “naughtier,” younger children focused on outcome rather than intention, and judged the person who caused more damage (even though it was accidental) to be naughtier. It was not until around the age of 10-11 years that children started evaluating on the basis of intentions (i.e., judging the intentional transgressor as naughtier than the accidental transgressor). Thus, Piaget’s work suggests that there is a “shift” from outcome-based to intention-based judgments across development.

However, from Piaget’s seminal work, it is unclear as to whether the shift from outcome-based to intention-based judgments is a result of an understanding of intention appearing later on in development, or whether it simply reflects a difference
in how children and adults weigh outcome and intention information. Thus, it is important to make the distinction between children being aware of mental states and the ability to use mental states in their evaluations (Imamoglu, 1975; Keasey, 1978). Although younger children in Piaget's task did not reliably use intention in their naughtiness assessments, this does not necessarily mean that they were unaware of the actors’ underlying mental states. Rather, the complexity of Piaget’s tasks may have potentially masked younger children's abilities to factor intention into their evaluations of others.

Firstly, Piaget’s task required children to hold two stories in their minds and they needed to compare the stories based on two competing variables (intentions and outcomes) before making their naughtiness assessments. Thus, the story-pair task may have been too cognitively taxing to adequately assess the abilities of younger children (see Arnsby, 1971; Berg-Cross, 1975; Wellman, Larkey & Somerville, 1979). Secondly, Piaget did not explicitly identify the actors’ intentions—children needed to infer them based on contextual information. In addition, the severity of the outcome was often confounded with intent: the accidental perpetrator’s act led to a much worse outcome than did the intentional perpetrator’s act. Thus, younger children may have focused more strongly on the amount of damage caused because it was the most salient cue in Piaget’s scenarios, rather than because children explicitly believed that outcomes were more important than intentions (see Karniol, 1978 for a review; see also Berg-Cross, 1975; Nelson, 1980; Nobes et al., 2009; Nunmedal & Bass, 1978).

Several follow-ups to Piaget’s seminal studies have supported the hypothesis that young children’s outcome-based evaluations were due to methodological
difficulties rather than a lack of intention understanding. When task demands were decreased by using single stories, actors’ intentions were highlighted, and/or outcomes were held constant, young children (as young as 3 years of age; Nelson, 1980) were able to use intention in their evaluations (see Baird & Astington, 2004; Karniol, 1978 for reviews; see also Berg-Cross, 1975; Keasey, 1978; Miller & McCann, 1979; Nummedal & Bass, 1976). However, these findings along with many others (e.g., Astington, 1986; Baird & Moses, 2001; Berndt & Berndt, 1975; Buchanan & Thompson, 1973; Farnill, 1974; Gutkin, 1972; Imamoğlu, 1975; King, 1971) still moderately support Piaget’s developmental shift claims in that they still find age-related increases in the use of intention in their judgments. When intention and outcome are directly pitted against one another (e.g., someone who has a positive intention but produces a negative outcome versus someone who has a negative intention but produces a positive outcome), younger children still make outcome-based judgments. Thus, although young children may be able to use intention judgments earlier than originally hypothesized in some situations, they still do not reliably privilege it until they are older.

However, more recent research in non-moral domains suggests that young children do show more sensitivity to actors’ underlying mental states when using implicit (or non-verbal) measures. For example, young children typically fail standard false belief tasks that involve verbal responses, but they pass these tasks when using measurements of predictive eye gaze or other non-verbal measures (e.g., Clements & Perner, 1994; Rubio-Fernández & Geurts, 2013). For example, Rubio-Fernández & Geurts (2013) told 3-year-old children modified false belief stories, using toy props. Children saw a girl putting her bananas into one of two fridges. When the girl was not
looking, the experimenter proceeded to move the bananas into the other fridge. When children were prompted to use the toys to continue acting out the story, they moved the girl towards the empty fridge, suggesting that they understood that the girl had a false belief that the bananas were in the fridge that she had originally put them in. Thus, non-verbal measures, provide additional support for children's implicit awareness of mental states, though children do not reliably use this information to make their explicit verbal responses. Using non-verbal paradigms, then, might also be a good way to determine whether infants are sensitive to mental states.

1.2 Infants' understanding of mental states

Indeed, research in non-moral domains suggests that infants have an early understanding of mental state concepts. Infants show sensitivity to beliefs (see Baillargeon, Scott, & Ze, 2010 for a review; see also Kovács, Téglás, & Endress, 2010; Senju, Southgate, Snape, Leonard, & Csibra, 2011) and intentions; infants interpret agents’ actions as goal-directed (see Woodward, 2005 for a review) and can infer unfulfilled (and thus unseen) goals, suggesting that their understanding of goals is mentalistic (e.g., Brandone & Wellman, 2009; Hamlin, Newman, & Wynn, 2009). Further, infants are more likely to imitate acts that are goal-directed and that are done intentionally rather than those that are not goal-directed or those that are done unintentionally (Carpenter, Akhtar, & Tomasello, 1998; Hamlin, Hallinan, & Woodward, 2008; Johnson, Booth, & O’Hearn, 2001; Lergestee & Markova, 2008; Meltzoff, 1995; Olineck & Poulin-Dubois, 2005; Olineck & Poulin-Dubois, 2009). Infants also display more negative affect and behaviors towards individuals who are intentionally mean.
rather than those who are accidentally mean (Behne, Carpenter, Call, & Tomasello, 2005; Dunfield & Kuhlmeier, 2010; Marsh, Stavropoulos, Nienhuis, & Legerstee, 2010).

In the classic ‘unwilling’ vs. ‘unable’ paradigms, infants interacted with an experimenter who either had a negative (and fulfilled) or positive (but unfulfilled) intention: she was either unwilling to share a toy with the infant (e.g., teased the infant with the toy) or she tried to share the toy with the infant, but was unable to do so (e.g., accidentally dropped the toy out of the infant’s reach). Infants as young as 6 months of age seemed to distinguish these interactions based on the experimenter’s intention; they were more likely to exhibit negative affect and gaze aversion when the experimenter was unwilling versus unable to share a toy (Behne et al., 2005; Marsh et al., 2010). In a follow-up study, Dunfield & Kuhlmeier (2010) examined how toddlers’ previous interactions with an ‘unwilling’ and an ‘unable’ actress would influence who they would choose to give a toy to in the future. Following the unwilling/unable task involving the two different actresses, toddlers watched as the actresses reached for a single toy that had fallen on the floor between them. Toddlers selectively gave it to the actress who had previously tried to share with them but had been unable to do so. Additionally, when another group of toddlers were given the choice to help an actress who had previously been successful in sharing a toy with them (fulfilled positive intention, good outcome) or an actress who had been unable to give them the toy (failed positive intention, bad outcome), toddlers helped the actresses equally. That is, toddlers’ decisions to help were primarily influenced by intention and not by outcome. Taken together, these studies suggest that intentions play a role in infants’ and
toddlers’ first-party evaluations of others and these evaluations subsequently influence toddlers’ own prosocial behaviors.

1.3 Infants’ third-party evaluations of morally-relevant actions

A growing body of research has started focusing on infants’ third-party evaluations, specifically as they apply to morally-relevant actions (e.g., Hamlin, in press; Hamlin, Mahajan, Liberman, & Wynn, 2013; Hamlin & Wynn, 2011; Hamlin & Wynn, 2012; Hamlin, Wynn, & Bloom, 2007; Hamlin, Wynn, & Bloom, 2010; Hamlin, Wynn, Bloom, & Mahajan, 2011; Schmidt & Sommerville, 2011). The benefit of looking at third-party evaluations is that infants are not directly affected by third-party interactions; they are neither benefited nor harmed by the interactions. As first-party actions are experienced first hand, they are presumably both highly salient and fairly easily interpreted. Thus, third-party evaluation tasks help to determine whether infants can go beyond evaluating those whose actions influence them directly—an ability that is clearly an essential requirement for moral evaluation.

The general design of infants’ third-party evaluation tasks involves infants watching two social events in alternation. In one event, a character behaves prosocially towards a Protagonist (e.g., helps the Protagonist achieve a goal), and in the other event, a different character behaves antisocially towards the Protagonist (e.g., hinders the Protagonist from achieving a goal). Following these events, infants are presented with the prosocial and antisocial characters and infants’ preference for one character over another is assessed using different preference methodologies depending on the age of the infant. For younger infants who are not yet capable of reaching, preference is
measured by examining infants’ relative visual attention to each of the two characters. Longer looking time towards one of the characters is used to indicate infants’ preference (Fantz, 1965). For older infants, a manual choice method is used; infants are asked to select the character that they prefer by reaching out and touching one of the two characters. Assessing infants’ preferences to approach (or alternatively to avoid) certain characters can indirectly inform us of their social evaluations of the characters, based on the assumption that they will approach those that they evaluate more positively and avoid those that they evaluate more negatively.

Findings from these tasks support the notion that infants as young as 3-months do differentially evaluate others, preferring those who act prosocially over those who act antisocially (e.g., Hamlin et al., 2007); infants prefer those who help others achieve a goal rather than those who hinder others from achieving their goals, and they also prefer those who give to others rather than take from them (e.g., Hamlin & Wynn, 2011). For example, Hamlin & Wynn, 2011 showed 3- and 5-month-old infants puppet shows in which a Protagonist puppet played with a ball. After tossing and catching the ball a few times, the Protagonist would drop the ball to one side of the stage. Then, either a Giver puppet would roll the ball back to the Protagonist, or a Taker puppet would take the ball and run away with it. After infants were habituated to these two events, their preferences for the Giver and the Taker were assessed. Three-month-old infants preferentially looked at the Giver puppet compared to the Taker puppet and 5-month-old infants preferentially reached for the Giver puppet. Additional work done by Hamlin et al., 2007 suggests that by 6 months of age, infants not only show an active
preference towards prosocial individuals, but also an aversion towards antisocial individuals.

Although these infant social evaluation tasks do not directly assess infants’ moral judgments per se (i.e., researchers cannot directly ask infants questions that are typically used to assess moral judgments in adults\(^1\); they cannot ask infants who is morally responsible, whether an action is right or wrong, blameworthy or praiseworthy, etc.), they do use measures that can perhaps indirectly inform us of their moral judgments. Presumably, infants’ assessments of whether agents are morally responsible should influence how they evaluate them (i.e. either positively or negatively), and these evaluations should subsequently affect their social preferences (i.e., who they decide to interact with). However, from these previous paradigms, it is difficult to discern whether infants are actually using *moral* assessments in their evaluations of others. Recall that mature moral judgments take into consideration the outcome of others’ actions, but they rely more heavily on the underlying mental states driving those actions. In the aforementioned paradigms, the prosocial and antisocial actors always have intentions that are consistent with the outcomes that they produce. For example, the *Giver* has a positive intention and acts according to that intention by giving the ball back to the Protagonist. By contrast, the *Taker* has a negative intention and acts consistently with that intention to produce a negative outcome (i.e. steals the ball). Thus, in these cases it is difficult to say whether infants are simply judging characters based on the observable outcomes that they produce (e.g., giving = good,

\(^1\) Although more recent research has assessed 19-month-old infants’ decisions to reward or punish individuals based on their past behaviors (Hamlin et al., 2011). After viewing prosocial and antisocial acts, infants preferred to reward prosocial actors (i.e. gave a treat to them) and punish antisocial actors (i.e. took a treat away from them).
taking = bad), or whether infants are privileging mental states in their evaluations of others (e.g., positive intent = good, negative intent = bad).

One way to disentangle whether infants are relying solely on observable outcomes or whether they are considering actors’ underlying mental states would be to pit outcome and intention against each other. Hamlin (in press) showed infants puppet shows involving two different events (depicted in Figure 1): (1) failed helping (2) failed hindering. During failed helping events, infants saw a Protagonist struggling to open a box to get a toy sitting inside of it. Then, a Failed Helper puppet attempted to help the Protagonist open the box, but was unable to do so. This resulted in the Protagonist failing to get the toy. During failed hindering events, infants again saw the Protagonist struggling to open the box to get the toy sitting inside of it. But this time, a Failed Hinderer puppet jumped on the lid, slamming the box shut. However, despite the Failed Hinderer’s attempts to block the Protagonist from getting the toy, the Protagonist still managed to open the box, fulfilling its goal. When infants were given a forced choice between the Failed Helper (who had a positive intention, but was associated with a negative outcome) and the Failed Hinderer (who had a negative intention, but was associated with a positive outcome), they preferentially reached for the Failed Helper by 8-months of age. These results suggest that infants can in fact evaluate moral agents on the basis of their underlying intentions, and they use those evaluations to guide their social preferences in future interactions.
Failed Helping

Protagonist tries to open box to obtain toy
Bear tries to help open box

Protagonist fails to open box, despite bear’s efforts to help
Protagonist fails to obtain toy

Failed Hindering

Protagonist tries to open box to obtain toy
Bear slams lid shut

Protagonist succeeds at opening box, despite bear’s attempts to hinder
Protagonist obtains toy

Figure 1. Failed helping and failed hindering events (Hamlin, in press)

1.4 Failed attempts vs. accidents

However, in the case of failed attempts an actor’s intentions are extremely salient. So even though the outcome achieved does not match the actor’s original
intention, their actions are still very clearly directed towards a particular goal: the content of their intention (i.e., to help open a box or slam a box closed) is directly readable from the actions themselves. In addition, failed actors are merely associated with a particular outcome, rather than causally connected to it. Thus, they may be less directly connected to the oppositely-valenced outcome than if they had directly caused it. In the case of accidents, however, these features are often reversed: actors’ intentions may be ambiguous or even absent when they accidentally do something, but they still play a causal role in producing a particular outcome. Thus, actors who accidentally do something are more directly linked to the outcome that they produce. For example, a driver who accidentally crashes into a pedestrian may not have intended to hit the pedestrian, but still played a causal role in injuring the pedestrian. Because actors’ intentions are more ambiguous in accidents than in failed attempts, and outcomes are more salient because they are causally connected to the actors, evaluating actors who cause accidents may be more difficult than judging actors who try, but fail to achieve a particular goal (e.g., Cushman et al., 2013; Hamlin, in press; Young et al., 2007). Indeed, children have a harder time exculpating accidental harmers and judge them more harshly than those who try, but fail to harm (e.g., Bärd & Astington, 2004; Cushman et al., 2013; Killen, Mulvey, Richardson, Jampol, & Woodward, 2011), and even adults assign some blame to accidental harmers in spite of them having innocent intentions (e.g., Cushman, 2008; Young et al., 2007; Young & Saxe, 2009).

Evaluations of accidental actors are further complicated by our perceptions of negligence (e.g., Heider, 1958; Nobes et al., 2009; Yuill & Perner, 1988). For instance, if we think that someone should have been able to anticipate a harmful outcome (even if
they did not intend for the harmful outcome to occur), we may blame them anyway to the extent that we perceive their actions to be negligent or reckless. In these cases, it is assumed that actors could have avoided these accidental harms by being more careful. By contrast, if we believe that there was no way that an actor could have anticipated the bad result, then we do not hold them morally responsible and we do not blame them for their actions.

Taken together, it is clear that evaluating moral agents is more difficult in the case of accidents than failed attempts. In the case of accidents, an actor's intentions are ambiguous or even absent, actors play a causal role in producing valenced outcomes (rather than simply being associated with them), and actors may be perceived as negligent. Thus, perhaps a more stringent test of whether infants are using moral assessments in their evaluations of others would be to see whether infants differentially evaluate intentional and accidental agents on the basis of their underlying mental states. Although previous research has shown that infants can distinguish between accidental and intentional acts by 14-months of age (e.g., Carpenter et al., 1998; Olineck & Poulin-Dubois, 2005), it is not yet known how infants will socially evaluate accidental and intentional agents in moral scenarios. Thus, the present investigation seeks to determine whether infants differentially evaluate intentional and accidental actors on the basis of their underlying mental states.
1.5 The present investigation

In the present study, 10-month-old infants viewed puppet shows in which a Protagonist was helped or hindered\(^2\) either intentionally or accidentally. Critically, across both the *Helping* and the *Hinder ing* conditions, the intentional acts were identical, as were the accidental acts. The only thing that made these actions prosocial or antisocial, then, was how they related to the Protagonist’s initial goal. That is, having the intention to help the Protagonist achieve a goal could be viewed more positively than not having any intention to help, whereas having the intention to hinder the Protagonist could be viewed more negatively than not having any intention to hinder.

Following the puppet shows, infants were given a forced choice between the intentional and accidental actors. If infants, like adults, privilege mental states in their evaluations of others, then they should subsequently prefer the actor who intentionally helps the Protagonist rather than the actor who accidentally helps, but they should prefer the actor who accidentally hinders the Protagonist rather than the actor who intentionally hinders.

Ten-month-old infants, in particular, were selected for this study because we know from previous research that infants privilege mental states when evaluating failed actors by 8-months of age (Hamlin, in press). Because evaluating accidental actors is thought to be more difficult than evaluating failed actors (e.g., Cushman et al., 2013; Hamlin, in press; Young et al., 2007), I decided to test a slightly older age group for the present study. Additionally, previous research shows that infants can distinguish

\(^2\) In prior studies (e.g., Hamlin et al., 2007) the term “hindered” is used to denote actions that prevent an agent from achieving a particular goal. However, the term “hindered” in the present study will refer to an action that results in a Protagonist’s goal being reversed.
between accidental and intentional actions by 14-months of age, so using a slightly younger age group can help us determine whether the ability to make this distinction emerges earlier than previously studied.
2 Experiment

2.1 Participants

Sixty healthy, full-term infants (32 male and 28 female, $X_{\text{age}} = 10$ months and 3 days, range = 9 months and 13 days to 10 months and 20 days) were recruited from the Greater Vancouver region. Forty-six additional infants were excluded from analyses due to experimental errors (13) fussiness (11) failure to choose either puppet (15), or parental interference (7).

2.2 Procedure

The experimental sessions were conducted in a 3.4 m x 3.7 m room. The room contained a black table (1.2 m x 1.5 m) that served as the puppet show stage. Three black curtains extended down from the ceiling, to surround the rear and two sides of the table. An additional curtain (approximately 60 cm from the rear curtain) was lowered between trials to occlude the stage. Infants sat on their parents’ laps at the front end of the table (approximately 60 cm from the occluding curtain), facing the puppet show stage. Parents were instructed to sit quietly for the duration of the study; they were asked not to communicate with their infants or to try to direct their attention in any way.

A puppeteer experimenter performed the puppet shows from behind the rear curtain. Her head and torso were completely hidden from the infants, and she wore a black long-sleeved shirt to hide her arms. A coder, who could not see the puppet shows and who was unaware of the puppets’ identities, recorded infants’ attention to each
event from behind the curtain on the infants’ right side; the coder observed the infants by peeking through a small hole in the curtain. Infants’ attention following each event was coded using the program JHab, and coding began when the puppeteer signaled to the coder that the action was over, by softly clicking her tongue. Infants’ attention to the stage was recorded until one of two conditions was met: (1) infants looked away for 2 consecutive seconds, or (2) a maximum of 30 seconds had elapsed.

All sessions were recorded using two video cameras: one camera was located in front of the rear curtain and was used to record the infants during the puppet shows and choice procedures; the second camera was located behind the infant and was used to record events on the stage. A digital mixer was used to combine the recordings from both cameras into one finalized video. Finalized videos consisted of a large main image of each infant’s face, with a smaller image of the stimuli embedded in the top left corner of the main image.

2.2.1 Familiarization

Infants were randomly assigned to either the Helping or the Hindering condition. In each condition, infants were familiarized to two Intentional and two Accidental events in alternation, for a total of 4 events. Events began when the occluding curtain was raised to reveal the stimuli. At the start of each event, infants saw two puppets standing on opposite sides of the stage, facing forward: (1) a leopard protagonist puppet and (2) one of two dog puppets: either the Intentional or the Accidental character. The dogs’ identities were distinguished by their shirt color (teal or purple), and each dog appeared on opposite sides of the stage. For example, if the Intentional character started out on the left side of the stage during intentional events (with the
Protagonist appearing on the right side of the stage), then the *Accidental* character would begin on the right side of the stage (with the Protagonist appearing on the left side of the stage) during accidental events.

A clear shelf made of Plexiglas (16 cm x 16 cm x 28.5 cm) was positioned between the two puppets on stage, and was always situated nearest to the dog puppet. A small yellow rubber duck sat on the floor of the stage, centered in front of the shelf. Additionally, a small stuffed moose either sat on top of the shelf (in the *Helping* condition) or beside the shelf, on the side nearest to the leopard puppet (in the *Hindering* condition).

In the *Helping* condition, infants saw events in which the Protagonist struggled and subsequently failed at getting the toy moose off of the shelf. The Protagonist’s goal was subsequently facilitated by each of the two dogs on alternating trials—both dogs knocked down the shelf, giving the leopard access to the toy. Critically, one dog pushed the shelf down intentionally, while the other dog accidentally bumped into the shelf and knocked it down. Thus, in both events, the Protagonist’s goal was facilitated, but only the *Intentional* character demonstrated a clear positive intention.

In the *Hindering* condition, infants saw events in which the Protagonist struggled, but eventually succeeded in putting a toy moose on top of the shelf. After this initial success, the Protagonist’s goal was subsequently hindered by each of the two dogs on alternating trials—both dogs reversed the Protagonist’s goal by knocking down the shelf, causing the toy moose to fall back down onto the floor. As in the *Helping* condition, one dog pushed the shelf down intentionally, while the other dog accidentally bumped into the shelf and knocked it down. Thus, in both events, the
Protagonist's goal was hindered, but only the *Intentional* character demonstrated a clear negative intention.

As previously noted, the *Intentional* and *Accidental* acts were the exact same across conditions. The *Intentional* acts always involved pushing down the shelf purposefully, while the *Accidental* acts always involved bumping into the shelf accidentally. However, what makes these acts prosocial or antisocial depends on what the leopard's initial goal was—knocking down the shelf after the leopard struggled to get the toy moose off of the shelf would be considered prosocial, whereas knocking down the shelf after the leopard had succeeded in putting the toy moose on top of the shelf would be considered antisocial. Each event is described in more detail below and an illustration of the critical actions in each event is depicted in Figure 2^3.

### 2.2.1.1 *Intentional helping*

The leopard and the dog turned simultaneously to “look”^4 at each other, and then turned back to face forward. The leopard ran over to the shelf, peeked around the shelf and looked at the toy duck. Then, the leopard looked up towards the top of the shelf, where the toy moose sat. With its arms outstretched, the leopard jumped up and down (twice) alongside the shelf, in an apparent attempt to reach the toy moose sitting on top of the shelf; this jumping action was repeated twice more, with the last attempt being slower and more exaggerated to emphasize the leopard’s struggle to reach the toy.

---

3 For the sake of brevity, not all actions were depicted in Figure 2. Only the critical actions relating to the Protagonist’s goal, or to the *Intentional* or *Accidental* actions performed by helpers or hinderers were included.

4 From this point forward, the use of the word “look” refers to how *we hoped* the puppets’ actions appeared to the infants.
Having been unsuccessful in retrieving the toy moose, the leopard lowered its head and ran to the back corner of the stage. The dog then turned to look at the leopard, and then up to look at the toy moose, in order to imply that the dog had inferred the leopard’s goal. The dog then ran directly to the side of the shelf and “intentionally” pushed it over, knocking the toy moose off of the shelf. To make the pushing action appear intentional, the dog puppet looked at the shelf, leaned into it, and used both paws to push it over. Once the shelf fell down and the toy was within reach, the leopard ran towards the shelf, picked up the toy moose, and bounced up and two times. This bouncing action was purposefully included to illustrate the leopard’s “happiness” at retrieving the toy moose. The dog then grasped the toy duck that had been resting on the floor in front of the shelf, and ran off stage.

2.2.1.2 Accidental helping

The leopard and the dog turned simultaneously to look at each other, and then turned back to face forward. Then, the dog ran off stage, disappearing from view and implying that the dog could not see the leopard’s goal-directed actions that followed. Once the dog had left the stage, the leopard ran over to the shelf, peeked around the shelf and looked at the toy duck. Then, the leopard looked up towards the top of the shelf, where the toy moose sat. With its arms outstretched, the leopard jumped up and down (twice) alongside the shelf, in an apparent attempt to reach the toy moose sitting on top of the shelf; this jumping action was repeated twice more, with the last attempt being slower and more exaggerated to emphasize the leopard’s struggle to reach the toy.
Having been unsuccessful in retrieving the toy moose, the leopard lowered its head and ran to the back corner of the stage. The dog then reappeared, turned to look at the leopard, and then turned to look at the toy duck sitting in front of the (clear) shelf. The dog ran directly towards the toy duck, in an apparent attempt to retrieve it. However, as the dog ran past the shelf, he clumsily bumped into it, knocking the shelf and the toy moose down. To increase the perception that knocking the shelf down was “accidental”, the dog continued to look at the toy duck (and not at the shelf) when the shelf was knocked over. Additionally, it was the side of the dog’s body (as opposed to his paws) that knocked the shelf down. Once the shelf fell down and the toy was within reach, the leopard ran towards the shelf, picked up the toy moose, and bounced up and two times. This bouncing action was purposefully included to illustrate the leopard’s “happiness” at retrieving the toy moose. The dog then ran off stage with the toy duck.

2.2.1.3 Intentional hindering

The leopard and the dog turned simultaneously to look at each other, and then turned back to face forward. The leopard ran forward and picked up the toy moose that was sitting beside the shelf. The leopard peeked around the shelf and looked at the toy duck, and then looked up towards the top of the shelf. Still holding the toy moose, the leopard jumped up and down (twice) alongside the shelf, in an apparent attempt to reach the top of the shelf; this action was repeated twice more, and on the final attempt, the leopard succeeded in reaching the top of the shelf. The leopard placed the toy moose on top of the shelf, and then jumped back down, landing to the side of the shelf.

To illustrate that he was “happy” about having placed the toy moose on top of the shelf, the leopard jumped up and down twice, and then ran to the back corner of the
stage. Following, the dog turned to look at the leopard, and then up to look at the toy moose, in order to imply that the dog had inferred the leopard’s goal. The dog ran directly to the side of the shelf and “intentionally” pushed it over, knocking the toy moose off of the shelf. To make the pushing action appear intentional, the dog puppet looked at the shelf, leaned into it, and used both paws to push it over. Once the shelf and the toy moose fell down, the leopard bounced up and down two times and placed its head down. This action was included to illustrate the leopard’s “disappointment” at his goal being hindered. The dog then grasped the toy duck that had been resting on the floor in front of the shelf, and ran off stage.

2.2.1.4  Accidental hindering

The leopard and the dog puppet turned simultaneously to look at each other, and then turned back to face forward. Then, the dog ran off stage, disappearing from view and implying that the dog could not see the leopard’s goal-directed actions that followed. Once the dog had left the stage, the leopard ran forward and picked up the toy moose that was sitting beside the shelf. The leopard peeked around the shelf and saw the toy duck. Then, the leopard looked up towards the top of the shelf. Still holding the toy moose, the leopard jumped up and down (twice) alongside the shelf, in an apparent attempt to reach the top of the shelf; this action was repeated twice more, and on the final attempt, the leopard succeeded in reaching the top of the shelf. The leopard placed the toy moose on top of the shelf, and then jumped back down, landing beside the shelf.

To illustrate that he was “happy” about having placed the toy moose on top of the shelf, the leopard jumped up and down twice, and then ran to the back corner of the stage. The dog then reappeared, turned to look at the leopard, and then at the toy duck
sitting in front of the (clear) shelf. The dog ran directly towards the toy duck, in an apparent attempt to retrieve it. However, as the dog ran past the shelf, he clumsily bumped into it, knocking the shelf and the toy moose down. To increase the perception that knocking the shelf down was “accidental”, the dog continued to look at the toy duck (and not at the shelf) when the shelf was knocked over. Additionally, it was the side of the dog’s body (as opposed to his paws) that knocked the shelf down. Once the shelf and the toy moose fell down, the leopard bounced up and down two times and placed its head down. This action was included to illustrate the leopard’s “disappointment” at his goal being hindered. The dog then ran off stage with the toy duck.

### 2.2.2 Choice

Following the familiarization period, parents were instructed to rotate their chairs 90 degrees clockwise, away from the puppet stage. Parents were asked to center their infants in their laps and to hold them firmly around their waist, so that infants were supported and could use both arms freely during the choice procedure. Parents were instructed to close their eyes as an experimenter (the online coder from the familiarization phase, who was blind to the puppets’ identities) emerged from behind the curtain, with the two puppets behind his/her back.

The experimenter knelt down in front of the infant and greeted him/her. Then, the experimenter said, “look” and presented the infant with the two puppets; the experimenter held the puppets approximately 30 cm apart and initially out of the infant’s reach. The experimenter ensured that the infant looked at each puppet before regaining the infant’s attention by pulling the two puppets back and saying “hi” to the infant. Once the infant looked back at the experimenter, the experimenter moved the
puppets forward such that they were now within the infant’s grasp. Infants were considered to have made a choice when they looked at and then touched one puppet, as determined online by the (blind) choice experimenter. The session ended after infants had either made a choice or 2 minutes had elapsed.

2.2.3 Counterbalancing and reliability

The shirt color of the intentional puppet (purple or teal), the order of events (intentional or accidental first), the side the intentional puppet appeared on stage during familiarization trials (left or right), and the side the intentional puppet appeared on during the choice procedure (left or right) were counterbalanced in each condition\(^5\). A second independent coder (also blind to the identities of the puppets) re-examined fifty-seven of the choice recordings for reliability. The coder could not examine the remaining three videos (1 Helping and 2 Hindering videos) as they were missing. The two coders reached 100% agreement.

\(^5\)This study was also run in conjunction with one other study that investigated infants’ sensitivity to generosity or greed in giving or taking scenarios. All of the same counterbalancing applied to this second study. Additionally, the order of presentation of the studies was counterbalanced across infants. There were no effects of order of presentation or valence of the other study (generosity or greed) on the results. Thus, all infants tested were included in the subsequent analyses.
Figure 2. Critical actions displayed in puppet show events. (A) Intentional helping events (B) Accidental Helping events (C) Intentional hindering events (D) Accidental Hindering events
2.3 Results

2.3.1 Looking time

A summary of looking times can be found in Table 1. Average looking times for each event type were calculated for each infant by adding infants’ looking time towards the first and second trials for each event type and dividing by two [e.g., (intentional trial 1 + intentional trial 2)/2]. A paired t-test was conducted to determine whether infants’ average looking time following intentional events differed from their average looking time following accidental events. Collapsed across both conditions, infants’ attention did not differ by event type \[t(59) = .08, p = .94, d = .01\]. Infants looked for an average of 7.67 seconds \((SEM = .69)\) following intentional events and 7.62 seconds \((SEM = .68)\) following accidental events.

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Intentional events</th>
<th>Accidental events</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First trial</td>
<td>Second trial</td>
</tr>
<tr>
<td>Helping</td>
<td>6.77</td>
<td>7.80</td>
</tr>
<tr>
<td>Hindering</td>
<td>9.26</td>
<td>6.87</td>
</tr>
</tbody>
</table>

Independent t-tests revealed that average looking time following intentional events did not differ by condition \([t(58) = .56, p = .58, d = .15]\) and average looking time following accidental events also did not differ by condition \([t(58) = .92, p = .36, d = .24]\). Infants in the Helping condition looked for an average of 7.29 seconds \((SEM = 1.01)\) following intentional events and 6.99 seconds \((SEM = .92)\) following accidental events. Infants’ attention did not differ by event type [paired t-test, \(t(29) = .31, p = .76, d = .06\)]. Infants in the Hindering condition looked for an average of 8.06 seconds \((SEM = .94)\).
following intentional events and 8.25 seconds ($SEM = 1.01$) following accidental events. Again, attention did not differ by event type [paired t-test, $t(29) = -.19$, $p = .85$, $d = .04$].

Interestingly, infants’ looking time did not differ towards first and second intentional trials [paired t-test, $t(29) = -.80$, $p = .43$, $d = .16$] or first and second accidental trials [paired t-test, $t(29) = .03$, $p = .98$, $d = .01$] in the Helping condition. However, infants’ looking time did marginally decrease between first and second intentional trials [paired t-test, $t(29) = 1.90$, $p = .07$, $d = .39$] and first and second accidental trials [paired t-test, $t(29) = 1.91$, $p = .07$, $d = .37$] in the Hindering condition.

2.3.2 Choice

Preliminary analyses revealed no effects of sex, shirt color of the intentional puppet, order that the intentional puppet was presented, or side that the intentional puppet was presented on during familiarization (all analyzed using Fisher’s exact tests, $p > .44$). However, there was a significant effect of side that the intentional puppet appeared on during choice on infant’s choice for the Hindering condition (Fisher’s exact test, $p < .05$) but not the Helping condition (Fisher’s exact test, $p = 1.0$). Even though the side that the intentional character appeared on during choice was almost evenly split within the Hindering condition (14 on the right, 16 on the left), infants were more likely to choose puppets appearing on their left-hand side compared to their right-hand side; infants chose the accidental character 14 times and the intentional character 7 times when it was on the infants’ left compared to 7 and 2 times respectively when the characters appeared on their right. However, because side preference during choice was not observed in the Helping condition or in previous studies using similar methods (e.g., Hamlin, in press; Hamlin & Wynn, 2011; Hamlin & Wynn, 2012; Hamlin et al.,
2007; Hamlin et al., 2010; Hamlin et al., 2011), we assumed that the side preference was not a reliable predictor of infant choice. Thus, in subsequent analyses, the data was collapsed across sex, shirt color of the intentional puppet, order that the intentional puppet was presented, and side that the intentional puppet was presented on during familiarization and during choice.

As predicted, infants’ preference for the intentional versus the accidental character differed by condition [$\chi^2(1, N = 60)= 11.28, p < .001, \phi = .43$], as depicted in Figure 3. Infants in the Helping condition preferred the intentional helper to the accidental helper (22 out of 30, Binomial, $p < .05$), while those in the Hindering condition preferred the accidental hinderer to the intentional hinderer (21 out of 30, Binomial, $p < .05$).

![Figure 3](image)

**Figure 3. Results of infants’ choice: Percentage of infants choosing the intentional and accidental characters in the Helping and the Hindering conditions**
The difference between infants’ total looking time following intentional and total looking time following accidental events was calculated to see whether longer looking to one event over another would correlate with infants’ choice (e.g., longer overall looking towards intentional events correlating with infants choosing the intentional puppet). Positive values indicated longer total looking time towards intentional events, while negative values indicated longer looking time towards accidental events. Although the mean difference scores were in the same direction as the choices made for each condition (Helping: $M_{\text{Int-Acc}} = .59$, Hindering: $M_{\text{Int-Acc}} = -.38$), no significant correlation between looking time and choice was found ($r = .12, p = .36$).

2.3.3 Excluded infants and additional analyses

Significantly more infants refused to participate in the Hindering condition than the Helping condition due to either fussiness (11) or failure to choose a puppet (15; there were 26 refusals in the entire study, and 21 of these infants were in the Hindering condition, Binomial, $p < .01$). Two coders, who were blind to condition and to the identities of the puppets, reviewed 9 (all in the Hindering condition) out of the 15 no choice videos to determine whether infants preferred to look at one puppet over the other during the choice procedure. The other videos (2 Helping and 4 Hindering) were omitted from analyses, due to both coders independently noting that they had difficulty viewing the infants’ eyes during the choice (e.g., infant leaned back and eyes were no longer in view). The two coders measured the total time infants spent looking towards the puppet on the infants’ left and right sides separately. The coders’ measurements for the total time infants spent looking at the intentional puppet (ICC = .98) as well as the
total time infants spent looking at the accidental puppet (ICC = .86) were reliably correlated

An average of Coder 1 and Coder 2’s measurements of infants’ looking time towards each puppet was calculated. A paired samples t-test revealed that infants’ attention towards each puppet did not differ \[t(7) = .39, p = .71, d = .15\]. Overall, infants’ looking time towards each puppet was quite short. During the 2-minute choice period, infants looked at the intentional puppet for an average of 9.8 seconds \((SEM = 3.30)\) and at the accidental puppet for an average of 8.6 seconds \((SEM = 2.76)\).

The high rate of refusals to participate in the \textit{Hindering} condition is suggestive that there is something particularly difficult or aversive about this condition for infants. To explore this idea further, additional analyses were conducted on the infants who completed the study. An additional coder, who was blind to condition and to the identities of the puppets, reviewed the recordings and coded the time it took infants in each condition to make their choice. Three out of sixty videos were missing (1 \textit{Helping} and 2 \textit{Hindering} videos), so these infants were excluded from subsequent analyses. A second blind coder analyzed 30\% of the videos for reliability; the two coders were highly reliable (ICC = .99).

Coding began as soon as the experimenter moved the puppets within the infant’s reach. If there was something inherently difficult or aversive about the \textit{Hindering} condition, then infants should have taken longer to choose between two \textit{Hinderers} than two \textit{Helpers}. However, the time that it took infants to choose a puppet did not differ by condition \[t(55) = .38, p = .71, d = .10\]. On average, infants in the \textit{Helping} condition took 15.86 seconds \((SEM = 3.74)\) to make a selection, while infants in the \textit{Hindering}
condition took 13.86 seconds ($SEM = 3.75$) to choose a puppet. Further, infants in each condition did not choose one character faster than another [$Helping$: $t(27) = .37, p = .72, d = 0.15$; $Hindering$: $t(26) = -.07, p = .95, d = .03$]. Infants in the $Helping$ condition who chose the intentional character, chose on average after 18.1 seconds ($SEM = 7.56$); those who chose the accidental characters, chose on average after 15.0 seconds ($SEM = 4.39$). Infants in the $Hindering$ condition who chose the intentional character, chose on average after 13.7 seconds ($SEM = 4.84$); those who chose the accidental characters, chose on average after 14.3 seconds ($SEM = 4.37$). Taken together, these analyses show that infants did not take longer to choose in the $Hindering$ condition than the $Helping$ condition, and they also did not differ in the amount of time it took to choose either the intentional or accidental character in each condition.
3 General Discussion

The results from the current study suggest that 10-month-old infants can make nuanced third-party evaluations of moral agents. Infants, like adults, evaluate others on the basis of their underlying mental states, and these evaluations subsequently influence their social preferences in future interactions. Infants viewed puppet shows involving helpers or hinderers and showed differential evaluations of actors depending on whether they acted intentionally or accidentally. Infants preferred those who intentionally helped a Protagonist achieve a goal rather than those who accidentally helped. By contrast, infants preferred those who accidentally hindered a Protagonist’s goal rather than those who intentionally did so. Notably, the opposing nature of the Protagonist’s goal in the two conditions (to get the toy off the shelf in the Helping condition, but to put the toy on top of the shelf in the Hinder condition) allowed the intentional and the accidental acts to be physically identical across conditions (i.e., intentional acts involved purposefully knocking over a shelf, while accidental acts involved inadvertently knocking the shelf down while in the process of trying to achieve some other goal), and all acts resulted in the same outcome (i.e., the toy moose falling off of the shelf). This suggests, then, that infants’ differential evaluations of helpers and hinderers were influenced primarily by their analyses of the actors’ mental states (intent vs. lack of intent) in reference to the Protagonist’s goal. These results cannot be attributed to differences in attention towards intentional or accidental events as infants’ attention to each event type did not differ. Additionally, when comparing
across conditions, infants did not differ in the amount of time spent looking towards either intentional or accidental events.

Although there were no overall differences in attention in either the *Helping* or the *Hindering* condition, there were several indices that suggest that there may have been something inherently aversive about the *Hindering* condition for infants. Firstly, infants’ attention did not diminish between the first and second intentional or first and second accidental trials for the *Helping* condition, but infants’ attention did marginally decrease in the second trial for each event type in the *Hindering* condition. One possibility is that infants may have been better at processing and remembering the negative events in the *Hindering* condition than the positive events in the *Helping* condition, and thus did not need to look as long to the second trials because they were “bored”. This interpretation would support the notion of a negativity bias in early development (see Vaish, Grossmann, & Woodward, 2008 for a review). Given this interpretation we might expect that infants in the past helper/hinderer studies should have habituated faster to hinderer events than to helper events. However, no mention of faster habituation rates for hindering than helping events was made in previous studies (e.g., Hamlin et al., 2007; Hamlin et al., 2010). Another interpretation of the results, however, might be that infants could have learned to anticipate what subsequent events would look like. After seeing the first couple of trials, infants in the *Helping* condition may have expected to see more helping events and thus may have continued to pay attention to these events because they were pleasing to them. By contrast, infants in the *Hindering* condition may have expected to see more hindering
events and thus may have paid less attention to subsequent events because they found them aversive and wanted to avoid looking at them.

Secondly, a significant proportion of the infants that were excluded due to fussiness or no choices were in the *Hindering* condition (21 out of 26 infants). This exclusion rate is much higher than previous studies involving one helper and one hinderer (e.g., Hamlin, in press; Hamlin, Mahajan, Liberman, & Wynn, 2013; Hamlin & Wynn, 2011; Hamlin & Wynn, 2012; Hamlin, Wynn, & Bloom, 2007; Hamlin, Wynn, & Bloom, 2010; Hamlin, Wynn, Bloom, & Mahajan, 2011). These past studies generally have had exclusion rates (due to fussiness or no choices) of less than 15%, whereas the *Hindering* condition in the present study reached 31%. The high exclusion rate found in the *Hindering* condition cannot be attributed to factors such as the banging noises caused by the shelf falling down scaring the infants, or the physical appearance of the puppets themselves being scary; these exact same features were present in the *Helping* condition, which had a similar exclusion rate (12%) to the previous studies involving one helper and one hinderer. Thus, observing four events involving a hindered goal (2 intentional and 2 accidental) and then having to choose between the two hinderers may have been particularly difficult for infants, whereas in previous studies, seeing helping events in alternation with hindering events may have been enough to offset infants’ negative affect to allow them to make it through the familiarization period and to make a choice between the two puppets.

Lastly, infants who made it through the familiarization period and refused to choose in the *Hindering* condition spent very little time (15% of the entire choice period) actually looking at the two puppets during the choice period. Out of the 120
seconds given to infants to make their choice, infants spent on average 9.8 seconds looking at the intentional puppet and 8.6 seconds looking at the accidental puppet. This might suggest that the infants found the puppets aversive and thus were avoiding eye contact with the puppets, although we cannot say this for certain as we were unable to compare this to infants who failed to choose in the Helping condition (due to their videos being unusable for coding).

In sum, infants’ decrease in attention towards second event trials, fussiness and refusal to choose, and gaze avoidance during choice in the Hinder condition suggests that there is something inherently difficult or aversive about the Hinder condition for infants. Although these events do seem to be aversive to infants, this does not seem to influence their ability to choose one character over the other for those infants who actually did make a choice. Infants in the Hinder condition did not take longer to make their decisions compared to infants in the Helping condition. Rather, infants chose equally as fast in both conditions, and there was no difference in the magnitude of the effect across condition, suggesting that the valence of the condition did not impact infants’ ability to make a choice.

Overall, the results from the present study in conjunction with Hamlin’s (in press) study examining infants’ evaluations of failed attempts, suggest that infants are quite capable of making nuanced evaluations of others on the basis of underlying mental states. Infants privilege intentions when evaluating failed attempts by 8-months of age, but even more impressively, they are able to differentially evaluate accidental and intentional actors (which is thought to be more difficult than evaluating failed attempts; e.g., Baird & Astington, 2004; Cushman et al., 2013; Hamlin, in press; Young et
al., 2007) on the basis of their underlying mental states by 10-months of age. The notion that infants can make these adult-like judgments contradict earlier claims of a shift from outcome-based judgments early on in development to intention-based judgments later on in development (e.g., Piaget 1932/1965, Kohlberg). Thus, young children’s failure to use mental states in their evaluations of others does not necessarily reflect a lack of awareness of mental states. Rather, the explicit nature (i.e., use of verbal responses) of the tasks themselves and the demand characteristics that they entail may mask children’s true ability to factor mental state information into their evaluations. Using a non-verbal paradigm in the present study allows us to tap into infants’ *implicit* awareness of the importance of mental states in their evaluations of others and shows that infants are able to make adult-like evaluations, preferring intentional helpers to accidental helpers, but accidental hinderers to intentional hinderers. This provides preliminary support to the claims that infants possess an early capacity to evaluate moral agents on the basis of their underlying mental states.

### 3.1 Limitations and future directions

The current study provides a first look into infants’ capacity to differentially evaluate intentional and accidental moral agents. However, there are several limitations to the current study that should be addressed in future work. First, although we know that infants prefer one character over the other, it is not clear as to whether infants are displaying preference because they view one character positively and the other negatively (or neutrally), or whether they perceive both characters to be positive or negative, but one is just more positive or more negative than the other. For example,
infants in the *Helping* condition may have judged both the intentional and accidental helpers to be prosocial (they both helped the Protagonist achieve his goal), but they may have evaluated the intentional helper more positively because he had the intent to help, whereas the accidental helper did not. Alternatively, infants may have viewed only the intentional helper as prosocial, while maintaining a neutral or even negative view of the accidental helper, because the accidental helper did not intend to help the Protagonist, but rather pursued his own selfish goal of retrieving a different toy. Similarly, infants in the *Hindering* condition may have judged both the intentional and the accidental hinderer as antisocial (they both hindered the Protagonist), but they may have judged the intentional hinderer more harshly because he had the intent to hinder, whereas the accidental hinderer did not. Alternatively, infants may have viewed only the intentional hinderer to be antisocial, while maintaining a neutral view of the accidental hinderer because the accidental hinderer did not have any knowledge of the leopard’s goal, and thus could not have intended to hinder the Protagonist. Thus, future studies should compare the intentional and accidental helpers and hinderers to a novel character that infants have no prior experience with. Based on prior research involving similar comparisons (e.g., Hamlin et al., 2007), infants who positively evaluate an agent should prefer the agent to the neutral character, while infants who negatively evaluate an agent should prefer the neutral character to the agent; infants who view an agent as neutral should be indifferent to the agent and the neutral character. Thus, by comparing the intentional and accidental helpers and hinderers to a neutral character, we can determine whether infants are displaying preference because they view one character positively and the other negatively (or neutrally), or whether they perceive both
characters to be positive or negative, but one is just more positive or more negative than the other.

Second, it is unclear as to what manipulations infants were responding to in the accidental events. The accidental acts in this preliminary study included two different manipulations to make it especially clear to infants that the accidental acts were unintentional: (1) the accidental character disappeared before the Protagonist displayed any goal-directed actions, and thus could not have known what the Protagonist's goal was (2) the accidental character did not intend to knock down the shelf, but rather bumped into it clumsily on his way to trying to fulfill his own separate goal. Thus, infants could have been responding to either or both of these cues in interpreting the actions as accidental. Thus, it is important to determine whether it is the accidental character’s ignorance or lack of intent that is influencing infants’ evaluations. To determine whether knowledge/ignorance plays a role, future work could examine whether infants differentiate between accidental agents that do or do not have knowledge of the Protagonist's goal (i.e., agent stays on stage during Protagonist's goal-directed actions or disappears before Protagonist exhibits any goal-directed actions) and also see how infants evaluate each of these agents compared to the intentional character. To determine whether an agent’s intent to knock down the shelf plays a role, future work should compare an agent who intends to knock down the shelf specifically to help or hinder the Protagonist, to an agent who intends to knock down the shelf, but as a means to achieving a goal of self-interest (e.g., perhaps there is a different toy on top of the shelf that the puppet is trying to retrieve, or perhaps the agent needs to knock down the shelf so that he can stand on it to get a toy that is even
higher out of reach). Although these comparisons are all useful to do, it is important to note that both knowledge and intent are *mental* processes. Thus, even though we do not know which of these two factors infants are responding to in the current study, we can at least be relatively certain that they are using *some* sort of mental state analysis in their evaluations of the actors.
4 Conclusions

The present study sheds light on infants’ ability to use mental state analyses in their evaluations of accidental and intentional moral agents. Infants prefer those who intentionally help rather than those who do so accidentally, but prefer those who accidentally hinder rather than those who do so intentionally. Infants’ abilities to make these adult-like judgments, points to the idea that this capacity to make evaluations of others on the basis of underlying mental states arises much earlier on in development than previously suggested.
References


