FEELING OUT THE ROLE OF FEELINGS IN INFANT SOCIO-MORAL EVALUATIONS

by

Conor Steckler

B.A., Western Washington University, 2009

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

Abstract

Research into infants' socio-moral evaluations has revealed that infants prefer prosocial to antisocial individuals, as demonstrated by their reaching behaviors (e.g., Hamlin, Wynn, & Bloom, 2007; Hamlin & Wynn, 2011). Although infants' choice behaviors have been demonstrated using several distinct social scenarios, the mechanism by which infants come to prefer one type of character to another is unknown. One possibility is that infants experience distinct emotions while observing prosocial and antisocial actions, and these emotional experiences guide their social preferences. As a first step in exploring this possibility, the current research used video-recordings of infants watching puppet shows with morally relevant content (prosocial and antisocial actions) and tested whether infants display more positive emotion towards prosocial acts and more negative emotion towards antisocial acts. Across three different studies and age groups, and two different methods, results provide support for the claim that infants' emotional displays differ when viewing prosocial versus antisocial acts.

Preface

This research was conducted with the permission of the University of British Columbia Office of Research Studies Behavioural Research Ethics Board, certificate number H10-01808.

Table of Contents

List of Figures

Acknowledgments

I would like to thank my advisor J. Kiley Hamlin for her support and guidance all along the way through the completion of this thesis. I would also like to thank my advisor Jessica L. Tracy and Mark Schaller for their very helpful comments on an earlier draft of this manuscript and for serving on my committee.

Chapter 1: Introduction

When we consider the acts of donating blood, stealing, or committing murder we easily recognize their moral significance. But by what kind of process do we come to view acts as moral, immoral, or amoral? That is, what are moral judgments fundamentally like and how did they get that way? It could be that mature moral judgments are made primarily through reasoning – a kind of cold-calculated process that operates in a qualitatively similar way to a number sense. From this perspective, a behavior, such as causing harm, is deemed "wrong" or "bad" similar to how it is wrong that 1+1=3. In contrast, moral feelings may be necessary for moral judgment. From this view, a behavior feels wrong or bad in a qualitatively different way than the recognition that $1+1\neq3$ and so is judged as bad/wrong. These are perspectives about what moral judgments might be like 'online' during the moment they are being made. But, to fully understand what moral judgments are like, we benefit greatly from studying development. If adults' (or older children's) moral judgments seem driven by emotion, a developmental period of reasoning about that moral content, entirely unemotional, may be necessary for and underlie those moral feelings. Alternatively, adults' moral judgments may require emotion at some point during development, but those judgments may be later expressed as non-emotional rules.

Increasingly since the Enlightenment, scholars have debated the importance of reasoning versus emotion for morality. Proponents of moral reasoning tended to be critics of an emotionally guided moral sense. For example, it was asserted that the only process by which one should or could conclude that something is morally right or wrong is via conscious reasoning, as it is only through reasoning that universal moral principles can exist. What did these scholars mean by "emotion" and why were they against it informing a moral sense? In general, emotion was (and is) a word used in reference to an internal force that influences thought and action, and

that arises spontaneously without conscious control (i.e., we cannot will an emotion into being the way we can will a thought into being).[1] The spontaneity of emotion was one reason why proponents of reasoning doubted its significance in the moral realm. For example, if emotion drove morality, then judgments of right and wrong may be swayed to and fro in accordance with uncontrollable fluctuations in feelings. From this perspective, how could universal moral laws exist if feelings differed from person to person and within the same person at different times? A well-known advocate of reasoning in morality was the 18th century philosopher Immanuel Kant. To Kant, a moral sense consisting of feelings or instincts could not yield universal moral principles. Instead, Kant thought that morality can and ought to be grounded purely in reason by the rational will, a force he saw humans uniquely possessing (1785/1959; see Denis, 2012 for a review of this philosophical debate). He wrote that:

> "[U]nless reason holds the reins of government in his own hands, a human being's feelings and inclinations play the master over him" (Kant 1797, 6:480)

In contrast to proponents of reasoning, to advocates of emotion, feelings were seen as something that motivated moral action and aided individuals in their determination of right and wrong. Champions of emotion in morality saw an otherwise insurmountable gap between "cold" reasoning processes and moral principles that feelings can account for. That is, if a moral principle is not *cared about*, then it is unlikely to be upheld as a principle at all. The 18th century philosopher David Hume advocated the view of an emotionally informed morality. Hume argued

[1] More recently, emotion has come to refer to a group of coordinated responses (e.g., facial expressions and action tendencies) occurring in the body and brain. These responses function to alter the cognitive, behavioral, or communicative state of the organism. Some aspects of an emotional response reach consciousness (e.g., subjective feelings) while other aspects are nonconscious (e.g., the release of a hormone that impacts the behavior of neurons; this definition draws heavily from Damasio, 1994; Nesse, 1990; Levenson, Ekman, & Friesen, 1990; see Russell, 2003 for a different view of emotion).

that morality required not only reasoning to discriminate moral content, but also sentiment or feeling to assign value to that content. In elucidating this, Hume wrote:

"Extinguish all the warm feelings and prepossessions in favour of virtue, and all disgust or aversion to vice: Render men totally indifferent towards their distinctions; and morality is no longer a practical study, nor has any tendency to regulate our lives and actions." (Hume 1777/1965, pp. 218-219)

1.1 Moral Judgment Versus Moral Action

Morality can be thought to occupy two related but distinct aspects of social life, that of moral judgment (wherein the actor evaluates another's actions and/or mental states as good/right or bad/wrong) and that of moral and altruistic behavior (wherein the actor behaves in ways that the actor and/or others consider good/right or bad/wrong, or in the case of altruism, behaves in ways that cost the self to benefit another). Although moral judgment and behavior are certainly linked, their relatively low positive correlation in individuals suggests that to some extent they are distinguishable and can be explored independently (e.g., Darley & Batson, 1973; Blasi, 1980). The current research focuses on the origins and nature of moral *judgment*, defined herein as the sense that some things are good/right while other things are bad/wrong, even when those things have no direct positive or negative value to the evaluator and when the evaluator engages in no moral action his or herself (i.e., even when the evaluator is a "3rd party"). Under this definition, moral rules are differentiated from mere conventions, or common modes of conduct (Turiel, 1983); nonetheless, the content of moral and conventional rules may differ across cultures and individuals (that is, what is moral for some may be conventional for others; e.g., Schweder, 1990).

1.2 Morality and Development

Stemming in part from the philosophical debate about morality, psychologists began a systematic investigation into moral development. Beginning with Piaget and expanded upon by Kohlberg (1969), research with children examined the development of moral reasoning. Kohlberg explored the cognitive development of morality by presenting his subjects with morally relevant dilemmas. For example, in his classic Heinz dilemma, a man called Heinz must decide whether to commit theft in order to save his dying wife. Kohlberg was interested not in the judgments his subjects came to (whether Heinz should or should not steal to save his wife), but rather in the reasoning process by which children arrived at their judgments. Overall, by examining the verbal reports from large groups of children, Kohlberg concluded that moral reasoning progressed in distinct stages characterized by unique calculations for determining right and wrong. For example, while young children might reason based on the rule "those who cause negative outcomes should be punished," older children might consider other factors in their calculations, such as whether the negative outcome was intended or accidental. Critically, Kohlberg characterized moral development as a process of qualitative change in the reasoning system, leaving emotional processes out.

1.2.1 Moral Development and Emotion

Compared to the role of reasoning, little is known about how emotion informs moral judgment across development. Hoffman (2000) argued that children are capable of experiencing "empathic anger" in response to a victim's distress caused by another. For example, the child may empathize with the victim and feel anger towards the perpetrator, and these emotions may then drive moral judgment. However, little empirical work has been done to explore whether such emotions do play a role in the child's judgment. Related to but distinct from moral

judgment, a number of studies have implicated the role of empathy in infants' and young children's morally-relevant (i.e., prosocial) *behaviors* (see Hepach, Vaish, & Tomasello, 2013 for a review). For example, 4 and 5 year olds' emotional reactions predict their prosocial behavior towards a distressed actor (Miller, Eisenberg, Fabes, Shell, 1996). Young children's sympathetic responses appear to be sophisticated in that they take contextual information into account: when a patient displays unjustified distress, young children sympathize with and help less than when the distress is justified (Hepach, Viash, & Tomasello, 2012). In studies such as Miller et al., although the actors display distress, the causes of the distress are often unrelated to the moral domain (e.g., a lone actor accidently falls from a swing and gets hurt), and the interest is on the child's prosocial *behaviors* (see also Warneken & Tomasello, 2006 for evidence of young children helping actors complete instrumental goals). Thus, although the child is motivated to help, it is difficult to assess from this literature that children, guided by emotion, think that they (and others) *ought* to help, or that it is wrong not to help. Although then that empathic responding may be related to moral judgment, this need not necessarily be the case. Thus, while evidence of the role of emotion on moral behavior may suggest that emotion also affects moral judgment, until this is empirically worked out it is difficult to say.

In contrast to research with young children, the relationship between emotion and moral judgment has received more attention in studies of adults (for a review see Prinz, 2006). Haidt (2001) claimed that many of humans' moral judgments are based on intuition, with reasoning often reflecting post-hoc rationalization processes, rather than being the driving force by which moral judgments are made in the first place. To test this, participants were presented with stories that were hypothesized to invoke moral judgments driven by intuitions. The stories were designed in such a way as to make it difficult for participants to come up with reasons for why

the transgression was morally wrong. In addition, the experimenter would play devil's advocate and combat participants' reasoning by pointing out why the specific reasons participants mentioned in support of their judgments were unjustified based on content of the stories. Subjects mostly condemned the acts in the stories (incest and cannibalism) and their moral judgments rarely shifted during the interview process. That is, confirming the role of intuition, moral judgments were maintained even after participants could no longer produce reasons to justify them. For example, in the case of incest, participants might default to reasoning that incest is wrong is because "it just is," or because "it is disgusting." Haidt called this phenomenon "moral dumfounding" (Haidt, Bjorklund, & Murphy, 2000; Haidt & Hersh, 2001).

A significant amount of research examining intuitive, emotional processes in adults' moral judgments has focused on the emotion of disgust. Across several studies and methods used to induce disgust, researchers have found that feeling disgusted leads to harsher judgments of moral violations (Wheatley & Haidt, 2005; Eskine, Kacinik, & Prinz, 2011; Horberg, Oveis, Keltner, & Cohen, 2009; Schnall, Haidt, Clore, & Jordan, 2008). For example, in one study, highly hypnotizable participants were told to feel disgusted in response to one of two arbitrary trigger words. After being brought out of a hypnotic state, participants rated moral transgressions that either included or did not include the trigger word. Participants were both more disgusted by and made harsher judges of moral transgressions in scenarios that contained the arbitrary trigger word compared to those that did not (Wheatley & Haidt, 2005). Less is known about the relationship between emotion and moral praise (i.e., the positive side of moral judgment; Haidt, 2003); one study suggests that witnessing prosocial actions leads to positive feelings of warmth in the perceiver and a desire to engage in prosocial acts, suggesting there is also a role for positive emotion in moral judgments (a feeling termed "elevation;" Algoe & Haidt, 2009).

Importantly, while these studies have found that emotion *intensifies* moral judgments, they have not ruled out that the judgments themselves do not arise independently of emotion (as pointed out by Huebner, Dwyer, & Hauser, 2009). While these studies on adults cannot provide direct evidence regarding the role of emotion in our earliest socio-moral judgments, they are at least suggestive of the possibility that emotion is important to an early developing socio-moral sense.

1.3 Emotion, Morality and Neuroscience

Other researchers have turned to neuroimaging to examine brain activity during moral judgment in its natural state, without any emotion induction techniques. Neuroscientific evidence involving normal adult participants suggests that some, but not all kinds of moral judgment and decision-making activate brain regions associated with emotion (e.g., Greene, Somerville, Nystrom, Darley, & Cohen, 2001; Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003). Greene et al (2001) described moral dilemmas to participants and then asked them whether certain actions were permissible or not. During the study, participants' brain activity was recorded with functional magnetic resonance imaging (fMRI). In one such dilemma, the trolley problem, there is a trolley speeding down a track on a path to kill five innocent people. The only way to save the five is to kill another person in their stead. In one version of this dilemma, by flipping a switch, the trolley can be diverted to another track on which an innocent person lies. Participants were asked whether it was permissible to flip the switch in this case. Most participants said that flipping the switch and killing the one person to save the five is permissible. In another version of the dilemma, instead of a switch, there is a bridge above the track and a passerby on top of the bridge. Participants were asked whether it is permissible to push the person off the bridge to save the five lives. Most participants said that the pushing is not permissible. Intriguingly, the outcomes in both scenarios are the same: one life is being sacrificed to save five. When it came

to responding to the case in which a person had to be pushed off the bridge, parts of the brain associated with emotion were more active. An increase in emotion-related brain activity did not occur in the case of flipping the switch. The conclusion Greene et al drew was that some kinds of acts (e.g., pushing) activate emotional processes that affect some of our moral judgments (whereas flipping the switch, because it is more impersonal, does not). This result suggests that emotion may be important in determining specific kinds of moral judgments but may not be recruited during all moral judgments.

1.4 Psychopathy and Moral Emotion

An additional clue about the importance of emotion for moral judgment comes from research on psychopathy. Psychopaths are characterized by their callousness and lack of empathy; often they are serious criminal offenders. Research indicates that psychopaths experience significantly reduced negative arousal in response to others' distress. This lack of arousal is selective: psychopaths do show increased arousal to threatening stimuli (Blair, 1997). Psychopaths are also less likely to see a distinction between moral and conventional rules (Blair, 1995). Thus, it may be the case that emotional responses toward others' distress are critically important for making moral judgments. This evidence should be interpreted somewhat cautiously however, given that there is no currently agreed upon explanation for why it is that psychopaths are capable of emotional reactions in some cases but not others (Blair, 2001). It may be that there is a deeper underlying (possibly computationally "cold") mechanism with which a lack of emotional responding to specific social phenomena co-varies.

1.5 What Might Emotion Function to Do?

Debates as to what the extant evidence reveals about the specific kind of relationship emotion has on our moral nature begs the larger question of what the functional role of emotion

on judgment and decision-making is thought to be. One influential perspective on emotion's role in cognition is referred to as "affect-as-information" theory, where affect refers to dimensions of positive and negative emotion more broadly. At its most fundamental core, affect-as-information suggests that we use our perceptions of our internal states (our feelings) to figure out what we think about things (including moral things). For example, on exceptionally sunny days (which presumably increase positive affect) participants report more life satisfaction (Schwarz & Clore, 1983; though see Lucas & Lawless, 2013 for contrary evidence), individuals may use the experience of guilt to repair transgressions (e.g., Ketelaar & Au, 2003) and low mood (particularly the associated fatigue) may be used to dissuade individuals from pursuing insurmountable goals (Keller & Nesse, 2005). The affect-as-information approach emphasizes research (e.g., Nisbett & Wilson, 1977) showing that many of the brain's computations occur outside of conscious awareness. Clore & Storbeck (2006), suggest that affect is the experience of conscious information about the nonconscious computations that gave rise to the affect (Haidt, 2001 made a similar point). Subjective feelings, then, arise in part from nonconscious computations and are used as guides in overt judgment and decision-making (Schwarz, 1990; Clore, Schwarz, & Conway, 1994; Damasio, 1994; though see Baumeister, Vohs, DeWall, & Zhang, 2007 for a competing view).

Other research indicates that the cause of affect or arousal can be misattributed and incorrectly used to guide judgments and decisions. For example, in one classic study, male participants interacted with a female experimenter on either a bridge that evoked feelings of arousal and fear or a bridge that did not evoke any emotion or arousal. Participants on the fear-arousing bridge were more likely to contact the experimenter after the study, presumably

attributing their arousal from fear of the height to arousal towards the experimenter (Dutton & Aron, 1974).

Misattribution effects may seem to lead to the conclusion that, because informational affect can lead to incorrect judgments and decisions, emotion is not a useful guide – presumably Kant would have argued for this. However, one compelling way to document the importance of affective information is to examine clinical populations with lesions in emotion-related parts of the brain. One well known patient, Elliot, suffered adult-onset damage to the ventromedial prefrontal cortex (Damasio, 1994). Elliot and other adult-onset patients with frontal damage studied by Damasio exhibited normally functioning moral reasoning (a point returned to below), social knowledge, executive function and working memory, and yet, he and other patients with similar brain damage showcased inappropriate social *behaviors* and had great difficulties in choosing between options even though they had the knowledge to discriminate between them.

Damasio hypothesized that these patients' difficulties were due to deficits in emotion-related parts of the brain, in particular, with using emotional information to inform future behaviors. To test this, Damasio and his colleagues designed a card sorting game to simulate decision-making in the real world (in which the rules are ambiguous and uncertainty exists). The card game consisted of four decks of two different types. Two of the decks gave high rewards but sometimes resulted in heavy losses. The other two decks tended to yield smaller rewards but less loss. The decks were rigged so that after some time the safer decks yielded more money. Participants played a game in which they were to earn as much play money as they could across an unknown number of trials (further adding to the ambiguity). Control participants (both undamaged and patients with non-frontal damage) learned over time which decks were better, presumably from information gleaned from previous wins and losses at each deck. Interestingly,

patients with frontal damage failed to learn which decks were financially better choices, and in fact gravitated towards the decks that were financially worse choices; they did not learn from previous wins and losses.

Further, this learning differential was specifically related to difficulties with emotion. In addition to tracking learning rates, all participants' skin conductance was measured.[2] Results suggested that both frontal patients and controls generated skin conductance responses to reward and punishment as they flipped over the respective cards generating these responses, suggesting they distinguished these outcomes on-line as they occurred. As the game continued on, control participants began generating anticipatory skin conductance responses before their choices, suggesting they were learning about the properties of the decks. In contrast, the frontal patients failed to generate anticipatory skin conductance responses, suggesting that they were unable to use prior emotional information (generated through wins and losses) to influence future decision-making toward or away from any particular option (Bechara, Damasio, Tranel, & Damasio, 1997; Damasio, 1994). What was especially intriguing about the frontal patients is that some of them *conceptually* distinguished which were the more advantageous decks toward the end of the game: they had the conscious knowledge of which decks paid out more in the long run. Yet, this knowledge did not inform their preferences; they continued to choose from the disadvantageous decks. This work suggests that discrimination and preference are dissociable, and that, normally, choice and preference rely in part on emotional information.

Although Damasio's frontal patients tended to engage in immoral behaviors and were impaired in the card sorting task, they were capable of moral reasoning (Saver & Damasio,

[2] Skin conductance is a measure of subtle increases in sweat on the skin brought on by the activity of the autonomic nervous system and is considered to be a measure of emotional reactivity.

1991). At first glance, these patients might support the claim that emotion is necessary for moral action, but unnecessary for moral judgment. Yet, as Damasio points out, it is possible that emotions play a critical role in moral judgments, *at some point during development*, but the judgments had already been internalized in the patients, who sustained their brain injuries in adulthood. That is, perhaps emotion was involved in the initial process by which the patients came to distinguish right from wrong, but is no longer needed. This empirical quandary may better be answered by examining our earliest socio-moral judgments, before emotion and reasoning processes feedback on one another and make a clear understanding of the role of each on moral judgment difficult to disentangle. As mentioned previously, developmental evidence in this domain is very sparse. The extant literature involved an investigation of two adults who both suffered damage to the frontal lobes before 16-months of age (Anderson, Bechara, Damasio, Tranel, & Damasio, 1999). The two patients were found to exhibit impairments in moral and non-moral decision-making (similar to adult-onset patients). In contrast to the adult-onset patients however, moral *reasoning* was also severely impaired. This result suggests that emotion may be critical to the normal development of the moral judgment system. However, these results should be interpreted as preliminary since the sample size was limited to two patients and the research was conducted several years after the onset of damage (suggesting that a host of experiences related to their deficits but distinct from emotion may account for their difficulties with moral reasoning).

Together, the reviewed findings provide valuable insights into the possible importance of emotions for moral judgments. They also, however, point to some critically important unanswered questions. In particular, the extent to which emotion, at some point during development, plays a causal role in moral judgment is mostly unknown. In addition, there has

been a lack of evidence examining any role, causal or not, of emotion's relationship to judgment across normal human development. Indeed, the little that is known about the early influence of emotion on judgment comes from clinical populations in which emotional deficits may co-vary with other deficits that are unknown, and in ways that are unknown. Thus, an exploration into the role of emotion in our earliest socio-moral judgments may provide valuable insight into the nature of those judgments and inform our understanding of the moral judgment system's typical development more generally.

1.6 Morality and Evolution

Why might humans possess a socio-moral sense that is developing in early ontogeny in the first place? In all studied human groups, there exist the concepts RIGHT and WRONG, although what is classified as right and wrong varies cross-culturally (Brown, 1991). The lack of variance in the existence of these two concepts across cultures suggests the possibility of an innate, naturally selected moral foundation. From an evolutionary perspective however, morality is somewhat mysterious. Most animal species seem to be amoral. Animals are designed via natural selection: the differential survival of competing "selfish" replicators (genes). Although "selfish" genes do not necessarily create selfish organisms, morality and altruism (i.e., suffering a cost for an unrelated individual's benefit) are nevertheless rare in species other than human beings (Dawkins, 1976; Alexander, 1984). How did humans acquire a moral sense, given its rarity in nature?

Multiple accounts of the evolution of morality and altruism exist and it is unlikely that any one of these can capture the complexities that shaped and continue to shape our moral sense (a point made by Nesse, 2009, see also Alexander, 1974; e.g., Hamilton, 1964; Trivers, 1971; Gintis, Henrich, Bowles, Boyd, & Fehr, 2007; Wilson & Wilson, 2007; Miller, 2007). For

example, one view suggests that distinguishing and preferring prosocial over antisocial individuals is so fundamental to successfully navigating the social world that this tendency was naturally selected (Williams, 1966, pp. 93-94). Because being praised versus shunned can have large survival and reproductive consequences, social preferences can influence the selection of traits such as prosociality and altruism (a term referred to as social selection; see Nesse, 2007). Thus, although moral behavior such as altruism may appear phenotypically costly, such behavior may be on average genetically beneficial.

1.7 Infants' Cognitive Capabilities

Consistent with the notion that a system or set of systems for moral judgment represents part of humans' naturally selected mental machinery, a growing body of literature suggests that young infants evaluate others based on their pro- and antisocial behaviors and the mental states underlying them (e.g., Hamlin, Wynn & Bloom, 2007; Hamlin & Wynn, 2011; Hamlin, in press).

Evidence of infants' socio-moral evaluations builds on several findings showcasing a variety of processes infants have been found to be sophisticated at. These processes include an understanding of agents and their goals (even when those goals are unfulfilled), an understanding of the physical world, and the possession of aspects of a Theory of Mind (the notion that others have mental states, which are unique to them, that cause their behaviors). In particular, infants have been found to distinguish agents from non-agents (entities with and without minds, respectively; Premack, 1990; Simion, Regolin, & Bulf, 2008). Infants have also been found to understand aspects of the physical world such as causality (Leslie & Keeble, 1987) and physics (Spelke, 1990). Other findings suggest that infants understand goal-directed action (Woodward, 1998; Woodward & Sommerville, 2000). Woodward (1998) habituated infants to an actor reaching for and grasping one of two toys. After habituation (i.e., a calculation of decreased

looking time suggesting adequate information processing) the toys switched locations. On test trials, the hand alternated between the same path with a new goal or the same goal but with a new path. If infants were encoding the actor's goal, they should not dishabituate to the new path the hand took, even though it looked physically different from habituation trials, since it would not be surprising to reach again for the preferred object. On the other hand, if infants were encoding just the physical act of reaching without an awareness of the agent's goal, then they should not dishabituate when the hand reaches in the same location, even though the object for which the hand is reaching has changed. On test trials, infants looked longer when the hand changed goals than when it changed paths. In a control condition in which the hand was replaced with an inanimate rod, there were no differences in looking time on test events. This suggests that infants were encoding the goal of the agent to which the hand belonged since they found it surprising when the goal, but not the physical appearance of the reaching changed.

Other research suggests that infants might operate on a "principle of rational action." According to the principle of rational action, infants expect agents to behave rationally (Csibra, Gergely, B´ıró, Koós, & Brockbank, 1999; Csibra, Biro, Koos, & Gergely, 2003). In Csibra et al (1999), infants were habituated to a shape on the right side of the screen reaching the left side of the screen by "jumping" over a barrier located toward the middle of the screen. On test events the barrier was removed and infants saw the shape either perform the rational (i.e., the most efficient) action to reach the other end of the stage or the same "jumping" action as before (i.e., the inefficient action). If infants do not expect movement based on rational action, they should be surprised at – and look longer towards – the movement in a straight line, as it is different than what infants were habituated to. If, however, infants expect movement to be rational, they should look longer at the jumping movement since the removal of the barrier renders jumping

unnecessary to reach the other side of the screen. Csibra et al found that infants were surprised at the identical movement when the barrier was removed and not by the novel movement, suggesting that infants expect movement to be rational (in this case, efficient).

In research providing evidence that infants can encode an agent's goal even when it is unfulfilled, Hamlin, Newman, & Wynn (2009) habituated 8-month-olds to some goal-relevant and some goal-irrelevant information. Specifically, infants saw an actor's hand attempt but fail to place a ring on a cone before moving away from the cone toward the ground and pausing (the goal-irrelevant motion). Although the goal was unfulfilled, some information was relevant to the goal (e.g., the cone was shaped such that it afforded the ring and the hand holding the ring move towards the cone) while other information (the movement toward the ground) was irrelevant to the goal. After habituation, infants saw two different test events, one consistent with the goal-irrelevant information of the hand moving away from the cone and the other consistent with the goal (the ring was successfully placed on the cone). Hamlin et al found that infants looked longer at the goal-irrelevant test trial in which the ring was placed on the ground compared to the goal-completion test trial in which the ring was placed on the cone. Infants' recovery from habituation on the goal-irrelevant test trial suggests that, during habituation, infants encoded that there was a goal despite its being unfulfilled. Two alternative explanations were ruled out in control conditions. In particular, infants did not show a baseline preference for looking at the ring on the ground compared to the ring on the cone (i.e., on just the test events without any habituation). Secondly, infants did not show differential looking on test trials after habituation when the hand was replaced with an inanimate claw (which other studies have found does not result in goal attribution by infants; e.g., Woodward, 1998; see also Brandone & Wellman, 2009).

Finally, a growing body of research suggests that infants are sensitive to the mental states of agents, including aspects of Theory of Mind such as agents' beliefs (Onishi & Baillargeon, 2005; Senju, Southgate, Snape, Leonard, & Csibra, 2011) and preferences (Kushnir, Xu, & Wellman, 2010).

1.8 Infant Socio-Moral Evaluations

Together, the reviewed evidence suggests that infants possess the building blocks necessary for socio-moral judgment involving the helping and hindering of others' goals. That is, infants possess an understanding of the physical world (such as that solid objects tend not to move through one another), they understand that agents can have goals, even when those goals are unfulfilled, and they understand that an agent has mental states that are unique to that agent. Much of the work examining infants' socio-moral judgments involves the helping and hindering of goals. Initial work depicted a 'Climber' character who repeatedly struggled to reach the top of a hill. One character, the "Helper," would bump the Climber up the hill thereby facilitating the Climber's goal. A different character, the "Hinderer," would knock the Climber down the hill, thereby blocking the Climber's goal (Kuhlmeier, Wynn, & Bloom, 2003; Hamlin, Wynn, & Bloom, 2007; see also Premack & Premack, 1997). In Hamlin et al (2007), after being habituated to these events in alternation, infants were presented with a choice between the Helper and Hinderer, with the assumption that infants reach for what they prefer. Infants (6- and 10-month-olds) were found to prefer the Helper to a Neutral character and a Neutral character to a Hinderer. Infants were at chance in their preferences however, when the Climber was replaced with a similar looking but inanimate shape that also lacked eyes (a cue to agency; Hood, Willen, & Driver, 1998). Infants' choice behaviors in Hamlin et al (2007) suggests that infants both positively evaluate (i.e., like) those who help others achieve their goals and negatively evaluate

(i.e., dislike) those who hinder others' goals. Since this early work, infants' preferences for helpers over hinderers has been replicated directly and conceptually, in infants as young as 3 months old (e.g., Hamlin, Wynn, & Bloom, 2010; Hamlin & Wynn, 2011). One such conceptual replication is giving and taking (Hamlin & Wynn, 2011). In this scenario, a Protagonist plays with a ball and accidently drops it in the direction of another character. One of two characters grabs the ball. The Protagonist proceeds to "ask" for the ball back by turning and facing the character clutching it. On their respective turns, one character, the Taker, runs off with the ball. Another character, the Giver, rolls the ball back to the Protagonist before leaving the stage. Here too, infants robustly preferred the prosocial over the antisocial character, but only when the Protagonist was cast as an agent: infants' preferences were random when the Protagonist was a mechanical, inanimate claw.

More recent research has examined more nuanced aspects of socio-moral judgment such as reward and punishment (Hamlin, Wynn, Bloom & Mahajan, 2011) and mentalistic versus outcome-based evaluation (Hamlin, Ullman, Tenenbaum, Goodman, & Baker, 2013; Hamlin, in press). Hamlin, Wynn, Bloom & Mahajan (2011) found that infants do not always prefer those who help others achieve their goals and avoid those who hinder others' goals. When the recipient of pro- and antisocial acts were *themselves* either previously prosocial or antisocial, an interesting pattern emerged. Specifically, 8-month-old infants were found to prefer the Helper (over the Hinderer) of a previously prosocial character, replicating the basic finding that infants prefer Helpers over Hinderers. In contrast, infants preferred a Hinderer (over a Helper) of a previously antisocial character. Importantly, the results were not driven by a preference for characters who matched the valence associated with the agent they acted upon. When the target of the helping and hindering was changed to the *victim* of a previous antisocial act, infants

preferred a Helper over a Hinderer (whereas the valence-matching hypothesis predicts that they would prefer the Hinderer). Valence-matching then, did not drive infants' preference for a character that hindered a previously antisocial character. In a third study, 19- to 23-month-olds were encouraged to either give a treat to a Helper or Hinderer (in one condition) or take a treat from a Helper or Hinderer (in a second condition). Building on infants' preferences for those who "reward" prosocial characters and "punish" antisocial characters, 19- to 23-month olds were found to themselves reward prosocial characters by giving them a treat and punish antisocial characters by taking a treat away. Together, these findings suggest developmental continuity between infants' preference for and action toward certain social characters.

In another study (Hamlin, in press), infants were found to evaluate characters based on their *intentions* to help or hinder. In particular, 8-month-olds saw and chose between different combinations of Successful and Failed characters who acted upon a Protagonist's goal to get a toy from a box. For example, one infant might see a Failed Helper attempt but fail to help the Protagonist get the toy from the box and, in alternation, also see a Failed Hinderer attempt but fail to stop the Protagonist from getting the toy from the box. In this case (as in other cases when intention was pitted against outcome) infants used intention information to inform their preferences, in this example choosing a Failed Helper over a Failed Hinderer.

Infants' "moral" sense does not seem to be confined to the facilitation and blocking of goals. Other research suggests that infants may possess an understanding of fairness (Sloane, Baillargeon, & Premack, 2012; Geraci & Surian, 2011). Geraci & Surian (2011) provided evidence that infants evaluate agents based on their tendencies to divide resources equally versus unequally, finding that infants prefer agents who distribute resources equally (but only to other agents; when the resource receivers were replaced with inanimate objects infants failed to show a

preference). Sloane et al (2012) found that infants' seemingly incorporate contextual information into their fairness expectations. Infants watched an experimenter ask two individuals to clean up a mess. After issuing the request, the experimenter left the scene. In a between subjects design, some infants saw one individual do all the clean-up work whereas other infants saw both individuals do some of the work. In front of the individuals were two boxes wherein, when applicable, the actors would put away what they had cleaned. For some babies, the box was transparent. For other babies the box was opaque. Once the mess was cleaned up, the experimenter returned and peered at each individual's box. All of the infants saw the experimenter divide stickers equally between the two individuals. Infants who previously saw only one of the individuals do all of the work looked longer when the experimenter divided the resources equally, but only when the experimenter could see into the box. Infants in the opaque-box control conditions looked similarly long whether one or both individuals had done the work, presumably because in the one-works opaque-box condition the experimenter lacked the knowledge that only one individual had worked.

Taken together, the reviewed findings on infants' capacity for socio-moral evaluation are intriguing for three reasons. First, in these studies, the infant is always a 3rd party. That is, infants are not participants or recipients of the characters' actions they later evaluate. This indicates that infants are doing something beyond responding to outcomes by which they are directly affected. Instead, infants are evaluating the (possibly novel) actions of novel characters upon other novel characters. Second, infants appear to be making judgments based upon agents' mentalistic properties (as in Hamlin et al., 2013). Lastly, infants have arguably not had many of the candidate experiences that would suggest that their capacities are (entirely) learned (as noted in

Hamlin, in prep). For example, it is unlikely that infants have experienced (or witnessed an agent) being pushed up or down a plane as in Hamlin, Wynn, and Bloom, 2007.

Together, the reviewed findings suggest that young infants possess a sophisticated computational system, or systems, for making socio-moral evaluations. Second, given that many of the studies reviewed were conducted on very young infants (sometimes as young as 3 months of age) some of our socio-moral architecture may be built-in or at the least easily acquired very early in development.

1.8.1 Using "Socio-Moral" To Describe Infants' Evaluations

Before proceeding, it is necessary to explain why infants' evaluations are described here as reflecting evidence of early socio-moral judgment, rather than social judgment more broadly. To begin with, infants' and toddlers' social evaluations and behaviors seem to map onto children's and adults' moral evaluations and behaviors. For example, 19- to 23-month-olds reward prosociality and punish antisociality in cases in which they were not personally involved. Reflective of mature moral judgments, infants' evaluations have been shown to take actors' mental states into account, rather than just the outcomes agents are associated with. Finally, infants' social preferences have been demonstrated in multiple distinct socio-moral domains. Thus, while it is recognized that there is great difficulty in concluding with certainty that infants' social evaluations are in any way moral evaluations, it seems, based upon the reviewed evidence, more likely than not that infants' evaluations reflect an early socio-moral sense.

1.9 The Present Research

It is, as of yet, entirely unclear *how* infants come to prefer one type of character to another. It is possible, in the Kantian sense, that infants possess socio-moral concepts and reason in ways that are entirely devoid of emotional processes. For example, infants may possess a basic

rule that facilitating goals = "good" or "right" just as infants possess concepts of number and recognize on some level that 1 + 1 = 2 is "right." Infants may similarly possess a basic rule that hindering goals = "bad"/"wrong" just as infants recognize on some level that $1 + 1 \neq 1$ (Wynn, 1992). A different possibility is that infants experience distinct *emotions* while considering moral and immoral content and subsequently use those emotional experiences to guide their social preferences. If it is true that infants have not had the candidate experiences to have learned about the social scenarios reviewed in the work on infants' evaluations, then infants' emotional responses to such scenarios may be similarly unlearned. Note that this view does not *pit* emotion against cognition. It is unlikely that emotional responses occur without computations to indicate whether an emotional response is warranted (such as computations of whether an agent's goal was helped or hindered). Rather, this model allows for such computations and predicts that emotion will be triggered at some point in the causal chain, leading to a response that motivates infants' preferences of social agents (that is, infants' emotional experiences will later drive their reaching behavior). This view is Humean in the sense that it suggests that *even if* infants can discriminate that one type of character is GOOD and another is BAD using computational rules (e.g., the oversimplified rule that agents that facilitate goals = good and agents that thwart goals = bad, if infants don't *experientially* react to such a distinction, they may be unlikely to make a judgment (i.e., show a strong social preference via their reaching behaviors).

It is in this light in which the current studies were carried out. Specifically, as a first step in examining the role of emotion in infants' socio-moral evaluations, the current studies tested the hypothesis that infants would display more positive emotion towards prosocial acts and more negative emotion towards antisocial acts. Importantly, a large body of research suggests that infants are capable of displaying positive and negative emotions. Ekman and Oster (1979) note

that the facial musculature is developed and usable by infants at birth and report evidence that almost all of the facial muscle actions used in adults' emotion expressions are identifiable by adults in newborns (Oster & Ekman, 1978). In particular, in terms of negative emotions, infants cry from birth (interestingly, also in response to other infants' crying noises; Sagi & Hoffman, 1971), display the facial expression of disgust (in response to certain tastes; Steiner, 1973; sometimes referred to as "distaste"; Rozin & Fallon, 1983), and show a startle response (Wolff, 1966). In terms of positive emotions, the capacity to smile is present in newborns although social smiling emerges around the 1st month of life (Wolff, 1963).

Chapter 2: Study 1

2.1 Method

2.1.1 Subjects The sample consisted of forty 22-month-olds from the Vancouver BC area (18 females). The mean age was 22-months 0 days (range 21 months 11 days to 22 months 26 days).

2.1.2 Procedure We examined the emotional reactions of toddlers' observations of prosocial and antisocial actions portrayed in puppet shows (see Figures 1 and 2 for pictorial representations of the shows). The shows consisted of 6 familiarization trials, with toddlers seeing 3 prosocial and 3 antisocial actions in alternation. The prosocial/antisocial events were giving a ball and taking a ball (as in Hamlin & Wynn, 2011). Each show begins with a ball resting in the middle of a stage surrounded by black curtains, with the Giver and Taker puppets sitting at each back corner. A Protagonist puppet enters from the center of the rear of the stage and runs up and grabs the ball. The Protagonist plays with the ball by jumping up and bouncing the ball up and down three times, after which he accidently drops the ball in the direction of either the Giver or the Taker. The Protagonist then "asks" for the ball back by turning towards the character who has the ball and opening his arms. The ball possessor then turns towards the Protagonist so that they are facing each other, and then both puppets turn forward simultaneously. This request/response combination repeats once; on the third request the shows differ depending on whether the ball possessor is the Giver or the Taker. During Giving events, the Giver rolls the ball back to the Protagonist and then runs offstage; the Protagonist catches the ball and freezes facing forward. During Taking events, the Taker runs offstage while still holding the ball. Once the Giver/Taker has left the stage, online familiarization coding begins. The trial ends and the curtain comes down and covers the stage when the baby looks away for 2 consecutive seconds, or after 30 seconds total. The following were counterbalanced across babies: The start order of the Giver

and Taker, the location of the Giver and Taker during familiarization events and the identity (i.e., the shirt color) of the Giver and Taker.

In order to code toddlers' emotional responses, all videos were trimmed to start just before the giving or taking took place and end 10 seconds after the action finished (or less if the infant looked away for 2 consecutive seconds before 10 seconds had elapsed). That is, for giving trials, coding began right when the ball was being given back to the Protagonist. For taking trials, coding began as the Taker was running off stage with the ball. We did not code the entirety of each trial, which could go as long as 30 seconds depending on when toddlers looked away from the display, as we reasoned that emotional reactions would be most apparent around the moments of giving and taking, and that emotional reactions after 10 seconds might reflect boredom or the effect of external distractions. To ensure experimenter blinding, trial order (1-6) and type (prosocial and antisocial) were randomized such that coders could not use any information from a previous trial to infer the event type of a subsequent trial, with the caveat that at least one of each event type was coded per toddler. Specifically, each video consisted of 4 randomized trials containing 3/1, 1/3, or 2/2 prosocial/antisocial events in random order. Sound was removed from videos to remove the possibility that coders could pick up on any auditory information that differed across prosocial and antisocial events; as was the face of the person (almost always a parent) whose lap the toddler was sitting on.

Before coding a baby's 4 trials, coders were instructed to watch the video through at least one time so as to become accustomed to the individual differences in the expressivity of the baby.

For each trial, four coders rated toddlers' dominant emotional expression on a scale from 1(*extreme negative affect*) to 4 (*neutral*) to 7 (*extreme positive affect*). This "naive coding"

method has been found to correlate highly with validated and more objective measures such as Baby FACS ($r > .90$; Oster, 2003), and has been utilized previously to reliably measure toddlers' emotional reactions to their giving to puppets (e.g., Aknin, Hamlin, & Dunn, 2012). Following the emotion ratings, the same four coders guessed which type of trial the toddler viewed (prosocial or antisocial) via a forced choice.

In addition to coding toddlers' general emotional expressivity, a second coder noted the extent to which these babies engaged in "eye-darting" (rapidly movements of the eyes) while viewing the puppet shows. Eye-darting (associated with anxiety and vigilance; Perkins, Inchley-Mort, Pickering, Corr, & Burgess, 2012) was coded on a scale from 0 (*no eye darting present*) to 1 (*slight to moderate eye darting*) to 2 (*extreme eye darting*).

To test whether toddlers' approach and avoidance behaviors differed by prosocial and antisocial trial type, one rater coded toddlers' forward and backward motions. Body movements were coded on a scale from -3 (*large body movement away from puppet show*) -2 (*small body movement away from puppet show*) -1 (*small appendage movement away from puppet show*) 0 (*no movement toward or away from puppet show*) 1 (*small appendage movement toward puppet show*) 2 (*small body movement toward puppet show*) to 3 (*large body movement toward puppet show*). Because toddlers have the tendency to move around frequently during the shows, multiple instances of movement were coded per trial if more than one forward or backward movement occurred. In addition, at the time of coding movement, the coder noted the position of the baby on the Z-axis (toward or away from the stage) on a scale from A (*leaned very much back from stage*) B (*leaned slightly back from stage*) C (e*rect/not leaning forward or backward*) D (*leaned slightly toward stage*) E (*leaned very much toward stage*). Movements were only coded when the baby was looking at the stage, which we took as an indication that the motion was more likely to

be relevant to the show than to other things in the toddler's surroundings (such as their caregiver or the door). Movements that did not seem purposeful were also not coded (e.g., when the caregiver adjusted their position and this adjustment caused the baby to lean a given way).

2.2 Results

The four coders reached adequate reliability, ICC =.80. All further analyses on emotion ratings and forced choice responding represent the averaged ratings among the four coders; however, eye-dart coding and body-movement coding represent results from a single coder, as only one coder rated the full dataset. Emotional expressivity among toddlers differed by trial type: $M_{prosocial} = 4.12$; $M_{antisocial} = 3.86$, $t(158)=2.68$, $p=.008$, two-tailed. Toddlers displayed more negative emotion while watching antisocial events, relative to the grand mean of their emotion expressions (4.0) used as an average or neutral point ($M_{antisocial} = 3.86$, one-sample $t(71)=1.88$, $p=.065$, two-tailed) and more positive emotion while watching prosocial events, relative to neutral ($M_{prosocial} = 4.12$, one-sample $t(87)=1.90$, $p=.059$, two-tailed). Tests against the neutral point of the scale yielded identical results, as the grand mean was the same as the neutral point of the scale.

Eye-darting did not differ by trial type ($M_{antisocial} = 1.07$, $M_{prosocial} = 1.01$, $t(154) = .60$, $p=.52$, two tailed). Body movements were coded such that, on any one trial there could be multiple coding entries. However, not enough trials were coded as having more than one entry, removing the possibility of meaningful statistical tests of the data across multiple coding instances in a given trial. Thus, the body movement data were analyzed using toddlers' first movement. No differences emerged in 22-month olds' tendencies to move toward or away from the stage on prosocial and antisocial trials, $M_{antisocial} = .20$, $M_{prosocial} = .34$, $t(154) = .91$, $p=.36$, two tailed. Other analyses were attempted that involved incorporating or controlling for infants'

body position at movement time 1, as well as incorporating those toddlers' trials in which more than one movement was coded through averaging their movements; all attempted analyses yielded null results.

Using just toddlers' emotional displays, coders were able to guess which type of trial the toddler was watching 58% of the time, statistically beyond chance levels (binomial test, p=.024, one-tailed). This statistic was computed by running a binomial test on the value arrived at from multiplying the averaged proportion correct across coders by the total number of trials (.5844 times 160 or 93 rounded down). This significant result suggests that, on average, coders can use toddlers' emotion displays to correctly inform their guesses about the type of trial being watched. Coders were equally likely to correctly guess the trial type infants were watching on prosocial and antisocial trials (59% of antisocial trials and 58% of prosocial trials, $t(158)$=.20, p=.80).

Previous research examining early socio-moral evaluations has shown that even young babies prefer prosocial to antisocial others. If emotion plays a role in social evaluations throughout infancy we should be able to see evidence of emotional responding in younger babies as well. Thus, in Study 2 we tested a sample of 7-month-old infants.

Chapter 3: Study 2

3.1 Method

3.1.1 Subjects The sample consisted of forty 7-month-old infants from the Vancouver BC area (20 females). The mean age was 7 months and 2 days (range 6 months 9 days -7 months 16 days).

3.1.2 Procedure The procedure was identical to that of Study 1.

3.2 Results

The four coders reached adequate reliability on ratings of emotion, ICC=.70. All further analyses on emotion ratings and forced choice responding represent the averaged ratings among the four coders (eye-dart coding and body-movement coding represent results from a single coder). Replicating the effect found with 22-month-olds, 7-month-olds' emotional expressivity was marginally higher for prosocial compared to antisocial trials, $t(158)=1.70$, $p=.09$, two-tailed. However, 7-month-olds' emotional expressions were not bi-directionally different from their grand mean of 4.08: $M_{antisocial} = 4.01$, one-sample against 4.08, $t(84)=1.17$, $p=.25$, two-tailed; $M_{prosocial} = 4.15$, one-sample against 4.08, $t(74)=1.24$, $p=.22$, two-tailed. A follow-up test against the neutral point of the scale revealed that 7-month-olds displayed more positive emotion on prosocial trials than neutral, one-sample $t(74)=2.59$, $p=.012$, two-tailed.

Suggestive of bi-directional emotional responding (positive emotion in the face, negative emotion in the eyes), 7-month olds were found to engage in more eye-darting in response to antisocial events than to prosocial events ($M_{antisocial} = 1.27$, $M_{prosocial} = 1.05$; $t(155.19)=2.33$, $p=.021$, two-tailed). As with Study 1, body movement coding did not reveal enough trials with more than one entry to be meaningful; thus, body movement data were analyzed using infants'

first movement. No differences between prosocial and antisocial trials emerged in 7-month olds' tendencies to move toward or away from the stage $t(158)=.35$, $p=.73$, two-tailed. Other analyses were attempted that involved incorporating or controlling for infants' body position at movement time 1, as well as incorporating those infants' trials in which more than one movement was coded through averaging their movements; all attempted analyses yielded null results.

Using just infants' emotional displays, coders were able to guess which type of trial an infant was watching 57% of the time, statistically beyond chance levels (binomial test, $p=.049$, one-tailed). This statistic was computed by running a binomial test on the value arrived at from multiplying the averaged proportion correct across coders by the total number of trials (.5734 times 160 or 91 rounded down). This significant result suggests that, on average, coders can use infants' emotion displays to correctly inform their guesses about the type of trial being watched beyond chance (50%). Statistically, coders were equally likely to correctly guess the trial type infants were watching on prosocial and antisocial trials (54% of antisocial trials correctly guessed versus 61% of prosocial trials; $t(158) = 1.38$, $p = .17$).

Studies 1 and 2 employed a randomization method, in which 4 out of 6 randomly selected trials were coded. Coders thus saw on average, and coded in succession, repeated trial types for a given baby. It is possible that coders implicitly picked up on the direction a baby was looking toward on each trial, and thus coders may have used their guesses on an early trial (e.g., where the baby was looking toward the left part of the stage) to inform their ratings on a later trial (e.g., where again the baby was looking toward the left part of the stage). To address this possibility, we sought to replicate the results with a pair method, described below.[3]

[3] Across coders, in only 15 out of 320 coding instances (in this case, a coding instance equals one baby's 4 trials) were incorrect forced choice guesses possibly systematic – meaning that only 4.7% of the time coders guessed all 4 trials incorrectly. If coders noticed the side of the show babies attended to, and if that information influenced their coding, then the systematic error rate should approach the total amount of errors. Instead, systematic errors are

Chapter 4: Study 3

4.1 Method

4.1.1 Subjects The sample consisted of sixteen 19-month-old infants from the Vancouver BC area (8 females). The mean age was 18-months and 29 days (range 18 months 17 days – 19 months 15 days).

4.1.2 Procedure We again examined the emotional reactions of infants' observations of prosocial and antisocial actions portrayed in the giving and taking puppet show. The shows consisted of 6 familiarization trials, with infants seeing 3 prosocial and 3 antisocial actions in alternation. The same 4 coders from Studies 1 and 2 coded all 6 of babies' familiarization trials. All further analyses on emotion ratings and forced choice responding represent the averaged ratings among the four coders. Videos were coded in local pairs such that trials 1 and 2 were coded one after the other, 3 and 4, and 5 and 6. Videos were randomized such that coders could see and code trial 2 before 1, or 1 before 2, etc. Further, across babies, pairs were coded in a randomized order such that coders did not code consecutive pairs from the same baby, to minimize the chance that coders could use information from prior coded pairs to inform later coded pairs. All videos were trimmed as in Studies 1 and 2, such that the researchers coded from the moment before the giving or taking took place until 10 seconds after the action finished (or less if the infant looked away for 2 consecutive seconds before 10 seconds elapsed). All other procedures to ensure experimenter blinding were identical to Studies 1 and 2. For each trial, four coders rated infants' dominant emotional expression on a scale from 1(*extreme negative affect*) to 4 (*neutral*) to 7 (*extreme positive affect*). Coders were instructed to watch a pair through at least once before coding, to become accustomed to the individual differences in the baby's

---

around 9 times less than total error. Nevertheless, attempting to replicate the results with an improved method is ideal.

expressivity. After rating the emotion of each trial in the pair, coders re-watched the pair and reported which trial they thought the baby was relatively happier watching (and thus which they thought was the prosocial trial). Coders also reported how confident they were in their forced choice on a scale from 0 (*complete guess*) 1 (*somewhat confident*) to 2 (*very confident*).

4.2 Results

One pair was excluded from analyses because the baby's face was not on camera for most of the 10 seconds (the coding results of this pair was in-line with predictions. Keeping or removing the pair did not change the interpretation of the results). The four coders reached adequate reliability, ICC =.86. All further analyses represent the averaged ratings among the four coders. The emotional displays of infants differed by trial type (prosocial or antisocial), $t(15)=3.22$, $p=.02$, two-tailed. Exploring the results through one-sample t-tests against the grand mean (of 4.23) yielded marginally significant bi-directional effects, $M_{antisocial} = 4.06$, one-sample $t(46)=1.81$, $p=.077$, two-tailed; $M_{prosocial} = 4.40$, one-sample $t(46)=1.57$, $p = .13$, two-tailed. A follow-up test against the neutral point of the scale revealed that 19-month-olds displayed more positive emotion on prosocial trials than neutral, one-sample $t(46)=3.63$, $p=.001$, two-tailed.

Using just infants' emotional displays, coders were again able to guess which type of trial the infant was watching 62% of the time, marginally beyond chance levels (binomial test, $p=.07$, one-tailed). This statistic was computed by running a binomial test on the value arrived at from multiplying the averaged proportion correct across coders by the total number *pairs* of trials (.615 times 47 or 29 rounded up from 28.91). This marginally significant result, notably with far fewer observations than Studies 1 and 2, is in line with the other binomial tests which suggest that, on average, coders can use infants' emotion displays to correctly inform their guesses about the type of trial being watched.

Chapter 5: Conclusion

Across three age groups and two different methods, infants' and toddlers' emotional displays differed by prosocial and antisocial trial type. Specifically, 7-month-olds displayed more positive emotion in reaction to prosocial events and engaged in more eye-darting in reaction to antisocial events. Nineteen-month-olds displayed more positive emotion in reaction to prosocial events but not more negative emotion in reaction to antisocial events. By 22-month-olds, the emotion reaction differences were bi-directionally different from neutral. That is, 22-month-olds displayed more positive emotion in reaction to prosocial trials relative to the average or neutral point, and more negative emotion in reaction to antisocial events relative to the average or neutral point. In contrast to 7-month-olds, 22-month-olds did not engage in more eye-darting in reaction to anti-social trials. That 22-month-olds' eye-darting did not differ by trial type but their negative facial expressions did suggests that perhaps the discrete emotional reactions infants experience towards antisocial events changes somewhere between 7 and 22 months.

Further, coders – completely blind to what type of show infants were watching, and hence using only infants' emotional displays – were consistently able to guess beyond chance levels whether babies were viewing prosocial or antisocial events. Together, these results point to the possibility that emotions play a role in our earliest socio-moral judgments. If this is the case, it suggests that not only may we be computationally equipped to understand the social world very early in ontogeny, we may be emotionally equipped to navigate it as well. Of note however, is that although infants have been found both to prefer prosocial characters to neutral characters and neutral characters to antisocial characters, the evidence of negative emotional reactions to antisocial events was less robust (at least in the facial musculature). It is possible that

infants *experience* but *do not display* negative emotion in response to antisocial acts reliably until later in development. Indeed, one of the only ways for infants to communicate their needs pre-verbally is through negative emotion. It is therefore possible that expressions of negative emotion in cases in which the infant is not signaling a need are inhibited, which may, speculatively, reduce the signal to noise ratio in their communication attempts. Nothing need be said about infants' conscious intentions to inhibit these responses, just that, if infants tended to display negative felt emotion in response to stimuli other than need signaling, this would increase the noise in their need signal. It is, however, a matter of speculation. The argument of inhibiting positive emotion signals, from a signal to noise ratio perspective, is not as strong (suggesting infants would not inhibit their displays of positive emotion). If one takes the view that positive emotion is displayed in the infant in part to shape the infant's environment by re-enforcing certain behaviors in the infant's caregiver, then there is little reason to suspect infants should refrain from displaying positive emotion at witnessing prosociality. A different possibility for infants' inconsistently displayed negative emotions toward antisocial actions is that young infants do not reliably *experience* negative emotion in their earliest moral judgments. This conclusion would suggest that negative emotion is not causally necessary for social evaluation of antisocial characters (though this conclusion is premature).

One question that may arise in readers' minds is whether infants' evaluations and the emotional responses correlated with viewing prosocial and antisocial actions are functional from the moments of their maturation. That is, the question may arise, why do mostly pre-verbal, pre-mobile infants seem to show such sophisticated evaluative capacities and seemingly feel similar states adults feel when witnessing socio-moral and socio-immoral acts? One answer is that the present results cannot rule out the possibility that infants' emotional reactions to 3rd party helping

and hindering are not learned responses. As an aside, it would interesting to test whether infants with siblings, who may be more likely to see 3rd party helping and hindering on part of their caregivers, display more valence-consistent emotion during the shows. Another answer is that, even if these emotional reactions represent some aspect of a built-in or early acquired emotional response to 3rd party interactions, not all that an organism comes equipped with or acquires early in development need be functional right away. Phenotypic features (including mental and behavioral tendencies) that are "built-in" must be built at some point in development, though they need not be functional until later in ontogeny (such as legs). From this view, what may be being tapped into in these studies is a kind of reflex to moral content that the infant does not actually use until later in life. Another possibility though, is that this system is functional, at least in a narrow sense. It may be the case that infants are actively learning, refining, and incorporating different systems together early in development such that they can adequately deal with the social world during childhood and beyond. Perhaps they are even incorporating a system that discriminates moral content with one that evaluates it emotionally. The tentative conclusion of these studies – that emotion plays a role in infants' socio-moral evaluations – comes with some caveats, however.

5.1 Limitations

First, emotional experience was measured by proxy of emotional display, and, as already discussed, experience and display need not align perfectly, making firm conclusions regarding how much display indicates experience in the infant somewhat unclear. Intuitively, emotional displays may possibly minimize the effect size of emotional experience between prosocial and antisocial trials since not all internal experiences lead to outward emotional displays. Research examining both emotional displays and autonomic activity among infants does indicate a

relationship between the two factors (Campos, Emde, Gaensbauer, & Henderson, 1975), though it is unclear how to extrapolate that to how well the effect size of emotion displays map onto the effect size of emotion experience.

Another limitation is that the studies were correlative in nature, so cannot preclude the possibility that emotion does not play a causal role in early moral judgments. As reviewed in the introduction, the extent to which emotion is a primary mediator of moral judgment is controversial. Some theorists of emotion argue that its phenomenological experience is too slow to act as a guide for judgment and action, suggesting instead that subjectively experienced emotion is more involved in learning from experiences that elicit emotion (Baumeister, Vohs, DeWall, & Zhang, 2007). Baumeister et al (2007) note for example that fears are often dealt with by unconscious processes in the brain before subjective feelings of fear arise. This suggests that (at least the phenomenological experience of) fear in such cases does not mediate the behavioral response (such as when leaping out of the way of a moving car and only later realizing what happened). However, the view that emotion mediates judgment and decision-making and the view that emotion plays a role in learning need not be orthogonal, and indeed the phenomenological experience of emotion is only one aspect of the multiple associated phenomena typically considered to be part of an emotion (e.g., Izard, 2010). In the case of the infant and toddler studies coded in the present paper, the distinction between judgment and learning from experience becomes somewhat murky. Infants and toddlers see repeats of the same trials and in the choice paradigm infants choose several seconds after the puppet show ends (suggesting that emotion-based learning and not judgment could theoretically play a role).

One further limitation is that only broad affect dimensions of positive and negative affect were considered, possibly masking subtle differences in the differential display of discrete

emotions such as anger or sadness. Also, the same socio-moral scenario was used across all three studies, the "ball show" (giving and taking; Hamlin & Wynn, 2011) so it is unclear at present whether infants display differential emotion to multiple kinds of prosocial and antisocial acts or whether the results are confined to giving and taking. Other scenarios were not used because they occasionally led infants to react in such a way as to make experimenter blinding improbable (e.g., blinking in response to a toy being dropped off a shelf), making this limitation difficult to overcome.

Lastly, it is possible that, because the physical actions of the prosocial and antisocial events differ, infants may be responding emotionally to the physical and not social differences between the two shows (although previous research has ruled out this possibility for infants' choice behaviors, it is nevertheless a possibility in the findings presented here; Hamlin & Wynn, 2011). While this is definitely a possibility, the fact that the evidence presented contained valence-matching emotional reactions to both positive and negative actions (to varying degrees across studies) suggests this explanation is somewhat less likely to be correct. That is, the physical difference account may be less parsimonious, requiring two explanations compared to the one social explanation – one for why infants displayed more positive emotion to prosocial events and one for why infants (tended) to display more negative emotion to antisocial events. The only physical account that requires one explanation is the possibility that infants' differential reactions were to the presence or absence of the ball (independent of why the ball was or was not on stage at the end of the show). Scores of studies have documented that infants' looking times may change in reaction to the disappearance of an object, though none has documented (though possibly because the researchers did not test for) differences in emotional reactions to an object's presence or absence (one study did note more 'aversion' in response to an object-permanence

violation infants did not seem to understand, though this was measured mostly by movement away from the stage. In the present studies, however, body movement coding yielded null results; Moore & Meltzoff, 1999). The result that 7-month-olds engaged in more eye-darting in reaction to antisocial versus prosocial trials may help tease the physical versus social explanations apart. As prior mentioned, eye-darting has been linked to the facial expression of anxiety (Perkins, Inchley-Mort, Pickering, Corr, & Burgess, 2012) and is associated with vigilance, scanning of the environment, and threat detection (Mathews, Mackintosh, & Fulcher, 1997). Further, anxiety has tentatively been linked to moral judgment in adults – individuals given anti-anxiety medication are more willing to condone 'personal' moral transgressions (as in the bridge/trolley dilemmas described above in Greene et al; but not 'impersonal' moral transgressions, suggesting the side-effect of sedation did not account for pattern of results; Perkins, Leonard, Weaver, Dalton, Mehta, Kumari, Williams, & Ettinger, 2012). The physical differences between the prosocial (giving) and antisocial (taking) scenarios might actually be expected to lead to more eye-darting in reaction to the prosocial scenario (in which the ball rolls across the stage back to the Protagonist). That 7-month-olds engaged in more eye-darting on antisocial trials relative to prosocial trials suggests then that the "physical-difference" alternative conclusion is less likely. Eye-darting on antisocial trials is also unlikely to arise from surprise: previous patterns of data from our lab suggest that infants do not look longer at antisocial actions versus prosocial actions (suggesting they do not find antisocial actions surprising). However, only future research can tease the physical versus social explanations apart in a more rigorous way.

5.2 Future Directions

One critical future step, as discussed above in the limitations, is to examine infants' emotional reactions to shows in which the same *physical actions* occur but are not social in nature (by, for example, replacing the animate Protagonist with an inanimate pincer; as in Hamlin & Wynn, 2011). This study is planned and will test the possibility that infants' emotional reactions reported herein may be driven by the physical – as opposed to the social differences between the two shows.

A further unanswered question is whether emotions play a causal role in guiding infants' moral evaluations. An experiment is planned that will test emotion's causal role in early moral judgment. The first will test whether emotions that are experimentally induced via manipulation of facial muscles alter infants' moral judgments. For example, if infants are induced to feel happiness, will they, likely adults, use that affect as information and become less likely to judge someone negatively? To induce emotion in infants, their facial muscles will be reconfigured with teething rings and other toys so that they form specific emotion expressions (such as a smile), which should then partially activate the corresponding emotional state (such as happiness); this method has been used previously to induce emotion and influence stimuli judgments in adults (e.g., Levenson, Ekman, & Friesen, 1990; Strack, Martin, & Stepper, 1988). Levenson et al (1990) found that participants who were instructed to manually configure their faces into different emotion expressions, muscle by muscle, experienced the emotion associated with the facial expression and showcased emotion specific patterns of autonomic nervous system activity. Strack et al (1988) found that having participants hold pencils in their mouths, such that the face formed a smile, influenced their positive evaluations of stimuli. In both of these studies, care was taken to rule out the possibility that participants were responding to demand characteristics in the

experimental setting, making this an unlikely explanation for the results. Rather, it suggests, as Levenson et al note, that there may be a link between the specific organization of facial muscles and specific patterns of autonomic nervous system activity associated with specific "affect programs"; and although the activation of these affect programs normally tend to *result* in the specific organization of facial muscles for the associated emotion, the program can actually be *triggered from* the activation of these facial muscles as well.

In particular, infants will view morally bad behaviours (again, taking, as in Hamlin & Wynn, 2011) while holding either a teething ring in their mouths to stimulate a smile, a soother to keep their facial muscles affectively neutral, or with nothing in the mouth. Taking an affect-as-information approach, it is predicted that infants' negative evaluation of the antisocial character (as measured through their relative choice between characters; e.g., Hamlin, Wynn, & Bloom, 2007) will be disrupted in the teething ring condition due to the positive emotion elicited by the teether. Critically, the soother condition will test whether such hypothesized disruption is due to infants' induction of emotion or to increased distractibility from entertaining a toy in the mouth during the show. It is predicted that emotion induction, and not increased distractibility, will alter infants' judgments. Findings in line with these predictions would suggest a *causal* link between emotion and socio-moral judgment in infancy and would support the Humean view of an early emerging, intuitive, emotionally driven moral judgment system.

One possibility is that these experiments may lack the power to find an effect. This research is new and exploratory and the intended manipulations may be subtler than is needed to produce the predicted effects. Specifically, it is possible that the manipulations are not sufficient to *reverse* or *block* infants' preferences, as the aforementioned study is designed to do. It may however, be more plausible that the aforementioned emotion manipulations will *intensify* infants'

preferences or nudge them in the valence-matched direction when infants are presented with ambiguity (e.g., in a case in which it is unclear whether a negative action was caused by an agent or not).

Thus, subtler tests have also been designed to be carried out alongside the more stringent tests. One tool that may be more sensitive than a choice measure is a looking time measure (due to the increased statistical power of continuous versus dichotomous measures). Previous findings by Hamlin, Wynn, & Bloom (2010) suggest that both 3-month-old and 6-month-old infants look longer at prosocial characters compared to antisocial characters in a side-by-side presentation, and that, for 6-month-olds this maps onto their reaching behavior (3-month-olds are unable to reliably reach for objects and so the choice method was not employed; McDonnell, 1975). Importantly, during the familiarization trials before the primary measure is collected, infants do not look longer at prosocial *scenarios* as they occur, compared to antisocial scenarios, suggesting that familiarity cannot account for their preferences. Building on this and testing the role of emotion in infants' preferences, infants will watch prosocial and antisocial events before being presented with the prosocial and antisocial character to look at. One group of infants will watch the show with valence-consistent emotion inductions whereas a second group of infants will watch the show with valence-inconsistent emotion inductions. Finally, a third group of infants will watch the show without any manipulation. Following this, infants' looking times will be compared between conditions, with the prediction that infants in the valence-matching condition will look longer at the prosocial character relative to the antisocial character compared to infants in the no-manipulation condition. In contrast, infants in the no-manipulation condition will look longer at the prosocial character relative to the antisocial character compared to infants in the inconsistent-valence manipulation condition. Findings in line with these predictions would

further support the role of positive and negative emotion in infants' evaluation of social others that engage in pro- and antisocial acts. Findings in line with the predictions in the first planned study may provide some of the strongest evidence for a causal role of emotion in socio-moral judgment. These studies will involve young infants and hence remove any possibility of demand characteristics present in many studies in adults that induce emotion. And further, study one will attempt to block emotion by inducing the valence-opposite, rather than merely showing that emotion intensifies socio-moral judgments that may well have already been made without emotion playing a primarily causal role.

Also of interest for future work is an examination of indicators of emotion other than the face, such as heart rate or skin conductance. Because there is some autonomic differentiation between both positive and negative emotions and *among* different positive and negative emotions, at least in adults (Levenson, Ekman, & Friesen, 1990), autonomic measures such as heart rate may help differentiate the discrete emotions infants experience during the puppet shows, thus allowing for a more fine-tuned assessment beyond broad affect dimensions (see also Campos, Emde, Gaensbauer, & Henderson, 1975). It would be interesting to see whether, as was found in the non-morally relevant card task among normal adult participants in Bechara et al 1996, infants, particularly young infants that typically need to be habituated over several trials, show increasing generation of anticipatory skin conductance responses before (or in general, to) antisocial actions across habituation, and whether infants that *do not* show such a response are more likely to choose the antisocial character.

One way to test Hume's and Damasio's discussed dissociation between moral knowledge and moral judgment in infants is to design a study in which infants' knowledge of a character may differ from their judgment of that character. There is one such study being conducted as of

this writing which may be showing this dissociation. Infants are shown three different puppet show scenarios. The first two serve as familiarization events and the last serves as test events. During the familiarization events, infants see a Helper character consistently help a Protagonist achieve their goal and a Hinderer character consistently prevent a Protagonist from achieving their goal. The show changes so that the Protagonist has a new goal, although the same Helper and Hinderer act in the same valenced way on the goal. Then, on the test events, one of the characters changes the valence of their behavior. In one condition, the prosocial action condition, both characters now help the Protagonist achieve a new goal. In this way, the former Helper before the test events performs the same valenced action whereas the former Hinderer switches the valence of their action and now helps the Protagonist. In the other condition, the antisocial action condition, on test events, both characters now hinder the Protagonist's new goal. In this way, the former Helper now hinders and the former Hinderer hinders again. If infants are discriminating between these characters and forming some kind of knowledge or impression about their dispositions, then they should look longer when, on the test events, a character performs a disposition-inconsistent behavior compared to one that does not. This is what this research has found so far. However, at the end of the study, infants are also presented with a choice between the two characters. Whereas the prediction might be that the infants should prefer the relatively more helpful character across the events (as opposed to a character who mostly hindered the Protagonist's goal then helped on tests events), the research thus far indicates that infants are at chance in their preferences between characters. This research suggests a possible dissociation between knowledge and preference (as in Bechara et al., 1996). However, unless there is a way to tap into infants' emotions while watching the shows, this possibility cannot be investigated (an alternative explanation is that maybe infants just got

confused). If this same experiment were to be repeated while infants' autonomic nervous system activity were being recorded, it may turn out to be the case that infants are not actually at chance in their preferences, but that some infants are able to use their emotional reactions to guide their preferences between the characters, whereas others are not, even though infants as a group conceptually distinguish between the characters.

The research presented herein provides evidence that infants display differential emotions when witnessing 3rd party prosocial and antisocial acts. The number of infants tested, puppet show trials coded, age ranges examined, coders used and different methods employed suggests that the captured phenomenon is robust across these factors. The results open the door for future studies to further elucidate emotion's earliest relationship to socio-moral judgment.

Figure 1. Antisocial Act

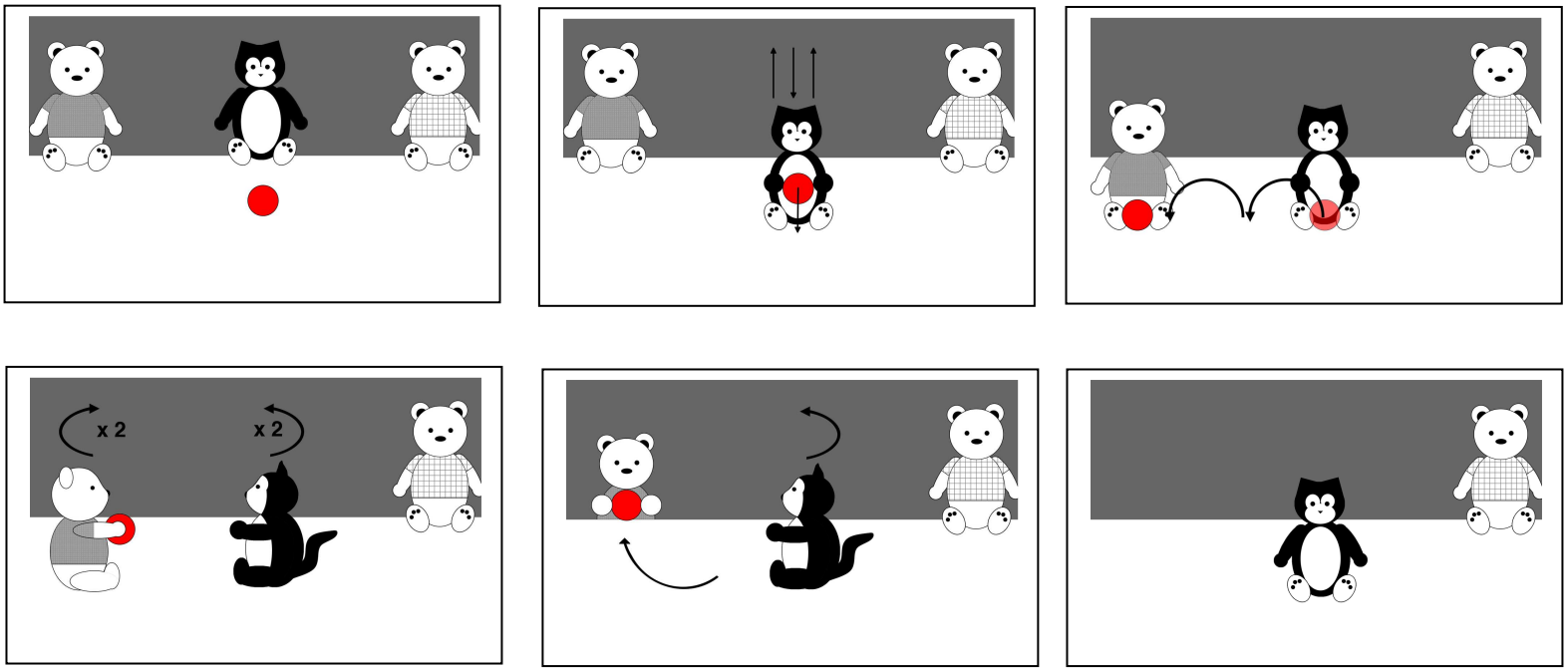


Figure 2. Prosocial Act

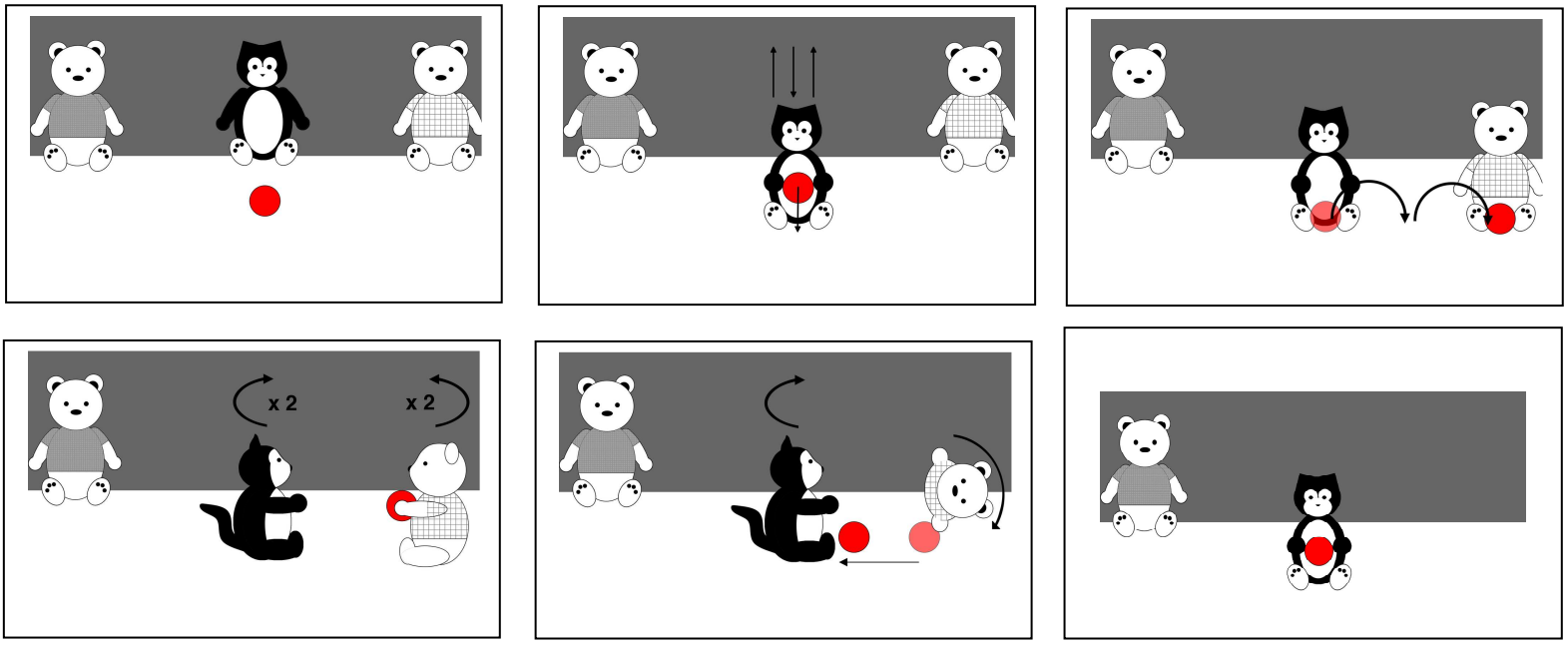
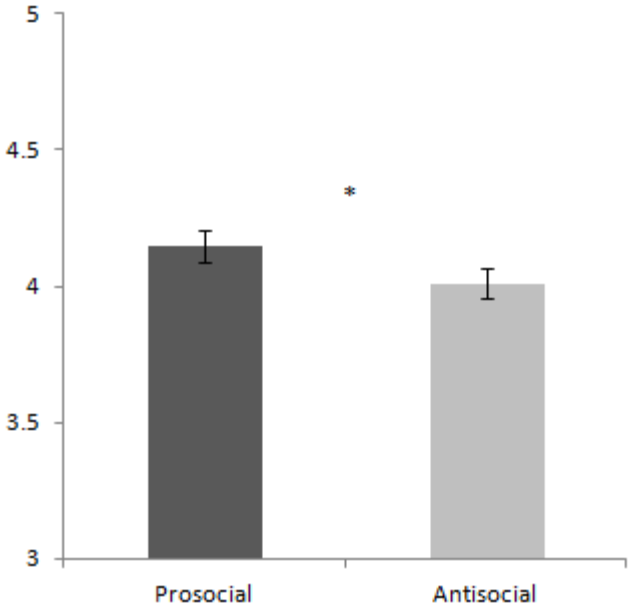
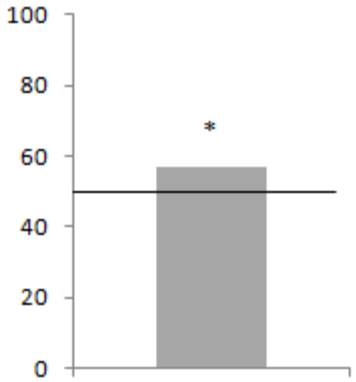


Figure 3. 7-Month-Old Emotion Display Results



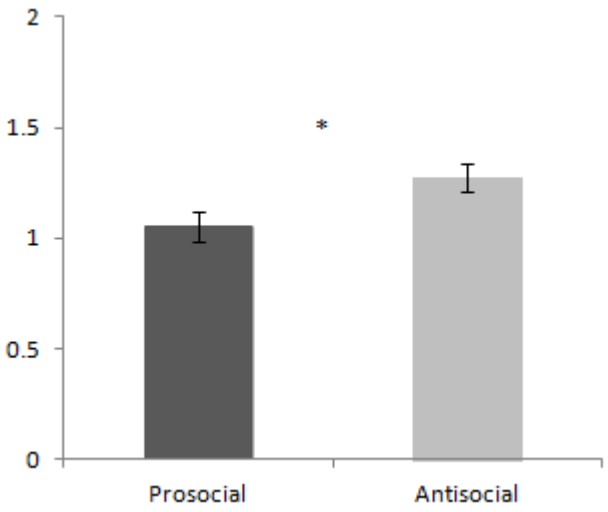
Averaged emotion ratings on the 1 (*extreme negative affect*) to 7 (*extreme positive affect*) scale. Note that the graph does not depict the full range of the scale. Error bars represent one standard error of the mean.

Figure 4. 7-Month-Old Forced Choice Results



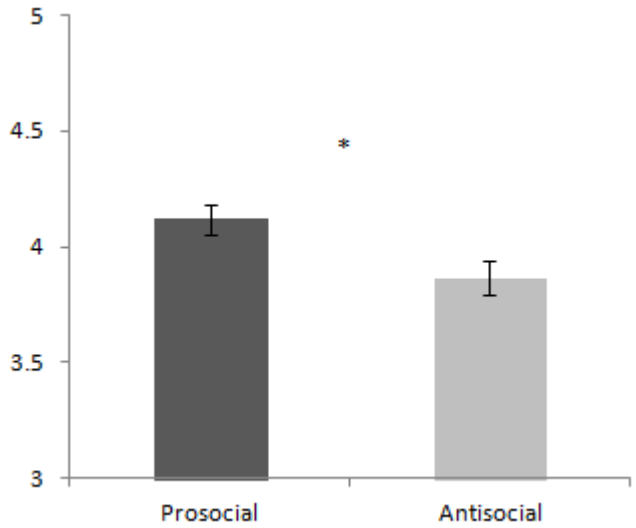
Averaged correct forced choice results. The dark line at 50% represents chance responding.

Figure 5. 7-Month-Old Eye-Darting Results



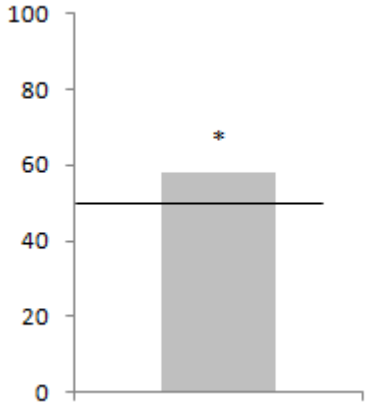
Higher numbers equals more eye-darting. Error bars represent one standard error of the mean.

Figure 6. 22-Month-Old Emotion Display Results



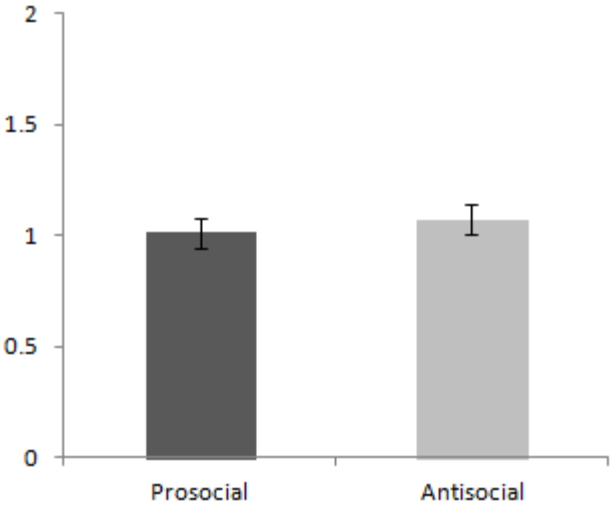
Averaged emotion ratings on the 1 (*extreme negative affect*) to 7 (*extreme positive affect*) scale. Note that the graph does not depict the full range of the scale. Error bars represent one standard error of the mean.

Figure 7. 22-Month-Old Forced Choice Results



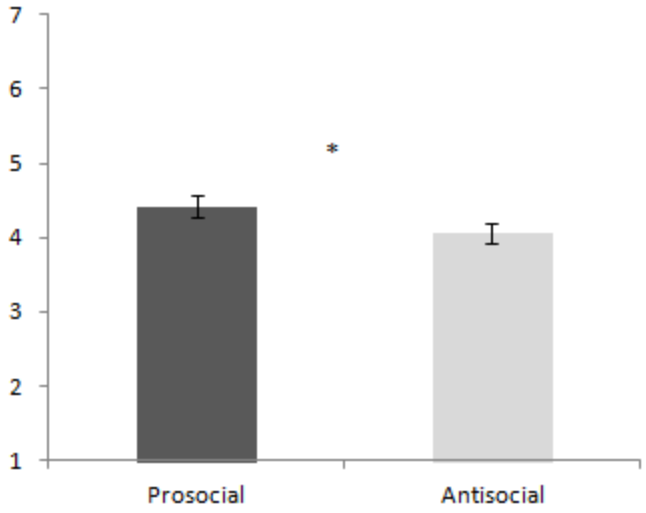
Averaged correct forced choice results.
The dark line at 50% represents chance

Figure 8. 22-Month-Old Eye-Darting Results



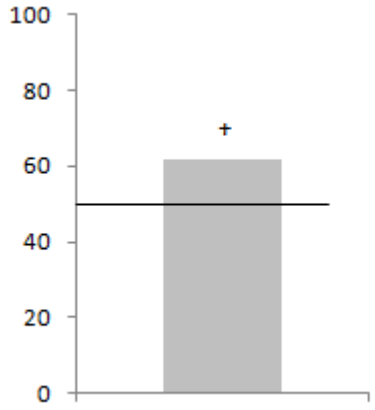
Higher numbers equals more eye-darting. Error bars represent one standard error of the mean.

Figure 9. 19-Month-Old Emotion Display Results



Averaged emotion ratings on the 1 (*extreme negative affect*) to 7 (*extreme positive affect*) scale. Error bars represent one standard error of the mean.

Figure 10. 19-Month-Old Forced Choice Results



Averaged correct forced choice results. The dark line at 50% represents chance responding.

References

Aknin, L. B., Hamlin, J. K., & Dunn, E. W. (2012). Giving leads to happiness in young children. *PLoS One*, *7*, e39211.

Alexander, R. D. (1974). The evolution of social behavior. *Annual Review of Ecology and Systematics*, *5*, 325-383.

Algoe, S. B. & Haidt, J. (2009). Witnessing excellence in action: the 'other-praising' emotions of elevation, gratitude, and admiration. *Journal of Positive Psychology, 4*, 105-127.

Baumeister, R. F., Vohs, K. D., DeWall, C. N., & Zhang, L. (2007). How emotion shapes behavior: Feedback, anticipation, and reflection, rather than direct causation. *Personality and Social Psychology Review*, *11*, 167-203.

Blair, R. J. R. (1995). A cognitive developmental approach to morality: investigating the psychopath. *Cognition, 57*, 1-29.

Blair, R. J. R. (2001). Neurocognitive models of aggression, the antisocial personality disorders, and psychopathy. *Journal of Neurology, Neurosurgery, & Psychiatry, 71*, 727-731.

Blair, R. J. R., Jones, L., Clark, F., & Smith, M. (1997). The psychopathic individual: a lack of responsiveness to distress cues? *Psychophysiology, 34*, 192-198.

Blair, R. J. R., Sellars, C., Strickland, I., Clark, F., Williams, A., Smith, M., & Jones, L. (1996). Theory of mind in the psychopath. *The Journal of Forensic Psychiatry, 7*, 15-25.

Blasi, A. (1980). Bridging moral cognition and moral action: a critical review of the literature. *Psychological Bulletin, 88*, 1-45.

Brandone, A., & Wellman, H. M. (2009). You can't always get what you want: Infants understand failed goal-directed actions. *Psychological Science, 20*, 85-91.

Brosnan, S. F., & De Waal, F. B. (2003). Monkeys reject unequal pay. *Nature*, *425*, 297-299.

Brown, D. E. (1991). *Human universals*. New York: McGraw-Hill.

Campos, J. J., Emde, R. N., Gaensbauer, T., & Henderson, C. (1975). Cardiac and behavioral interrelationships in the reactions of infants to strangers. *Developmental Psychology, 11*, 589-601.

Csibra, G., Bıró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, *27*, 111-133.

Csibra, G., Gergely, G., Bíró, S., Koos, O., & Brockbank, M. (1999). Goal attribution without

agency cues: the perception of 'pure reason' in infancy. *Cognition*, *72*, 237-267.

Clore, G. L., Schwarz, N., & Conway, M. (1994). Affective causes and consequences of social information processing. *Handbook of Social Cognition*, *1*, 323-417.

Damasio, A. R. (1994). Descartes' error: Emotion, rationality and the human brain. *New York: Putnam*, *352*.

Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L.B.L., Parvizi, J., & Hichwa, R. D. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience, 3*, 1049-1056.

Darley, J. M., & Batson, D. (1973). "From Jerusalem to Jericho": a study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology, 27*, 100-108.

Dawkins, R. (1976). *The selfish gene*. Oxford University Press.

Denis, Lara. "Kant and Hume on Morality", *The Stanford Encyclopedia of Philosophy* (Fall 2012 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2012/entries/kant-hume-morality/>.

Ekman, P., & Oster, H. (1979). Facial expressions of emotion. *Annual Review of Psychology*, *30*, 527-554.

Eskine, K. J., Kacinik, N. A., & Prinz, J. J. (2011). A bad taste in the mouth: Gustatory disgust influences moral judgment. *Psychological Science*, *22*, 295-299.

Geraci, A., & Surian, L. (2011). The developmental roots of fairness: infants' reactions to equal and unequal distributions of resources. *Developmental Science, 14*, 1012-1020.

Gintis, H., Henrich, J., Bowles, S., Boyd, R., & Fehr, E. (2008). Strong reciprocity and the roots of human morality. *Social Justice Research*, *21*, 241-253.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293,* 2105-2108.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review, 108*, 814-834.

Haidt, J., Bjorklund, F., & Murphy, S. (2000). Moral dumbfounding: When intuition finds no reason. (Unpublished manuscript, University of Virginia).

Haidt, J., & Hersh, M. A. (2001). Sexual morality: The cultures and emotions of conservatives and liberals. *Journal of Applied Social Psychology*, *31*, 191-221.

Haidt, J. (2003). Elevation and the positive psychology of morality.

Hamilton, W. D. (1964). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, *7*, 1-16.

Hamlin, J. K. (in prep). Does the infant possess a moral concept? *University of British Columbia*

Hamlin, J. K. (in press). Failed attempts. *University of British Columbia*

Hamlin, J.K., Newman, G., & Wynn, K. (2009). Eight-month-old infants infer unfulfilled goals, despite ambiguous physical evidence. *Infancy, 14*, 579–590.

Hamlin, J.K., & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitve Development, 26,* 30-39.

Hamlin, J.K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature, 450,* 557-559.

Hamlin, J. K., Wynn, K., & Bloom, P. (2010). Three-month-olds show a negativity bias in their social evaluations. *Developmental Science, 13,* 923–929.

Hamlin, J.K., Wynn, K., Bloom, P., & Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proc Natl Acad Sci USA.* doi: 10.1073/pnas.1110306108.

Hamlin, J. K., Ullman, T., Tenenbaum, J., Goodman, N., & Baker, C. (2013). The mentalistic basis of core social cognition: experiments in preverbal infants and a computational model. *Developmental science*, *16*, 209-226.

Hepach, R., Amrisha, V., & Tomasello, M. (2013). A new look at children's prosocial motivation. *Infancy, 18*, 1-24.

Hepach, R. Amrisha, V., & Tomasello, M. (2012). Young children sympathize less in response to unjustified emotional distress. *Developmental Psychology,* Advance online publication. doi: 10.1037/a0029501

Hoffman, M. L. (1983). Affective and cognitive processes in moral internalization: An information processing approach. In E.T. Higgins, D. Ruble, & W. Hartup (Eds.), *Social Cognition and Social Development: A Socio-Cultural Perspective*. New York: Cambridge University Press, 236-274.

Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice.*

Hood, B. M., Willen, D. J., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science, 9*, 131-134.

Horberg, E. J., Oveis, C., Keltner, D., & Cohen, A. B. (2009). Disgust and the moralization of purity. *Journal of Personality and Social Psychology*, *97*, 963-976.

Huebner, B., Dwyer, S., & Hauser, M. (2009). The role of emotion in moral psychology. *Trends in Cognitive Sciences*, *13*, 1-6.

Hume, D. (1777). 1965. *An enquiry concerning the principles of morals.*

Izard, C. E. (2010). The many meanings/aspects of emotion: definitions, functions, activation, and regulation. *Emotion Review, 2*, 363-370.

Kant, I. (1785). 1959. *Foundation of the metaphysics of morals*.

Keller, M. C., & Nesse, R. M. (2005). Is low mood an adaptation? Evidence for subtypes with symptoms that match precipitants. *Journal of Affective Disorders, 86*, 27-35.

Ketelaar & Au (2003). The effects of guilty feelings on the behavior of uncooperative individuals in repeated social bargaining games: An affect-as-information interpretation of the role of emotion in social interaction *Cognition and Emotion, 17*, 429-453.

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In *The Handbook of Socialization Theory and Research,* D.A. Goslin, (Ed.). Chicago: Rand McNally.

Kuhlmeier, V., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological Science*, *14*, 402-408.

Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science, 21*, 1134-1140.

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition, 25*, 265-288.

Levenson, R. W., Ekman, P., & Friesen, W. V. (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology*, *27*, 363-384.

Lucas, R. E., & Lawless, N. M. (2013). Does life seem better on a sunny day? Examining the association between weather conditions and life satisfaction judgments. *Journal of Personality and Social Psychology, 104*, 872-884.

Mathews, A., Mackintosh, B., & Fulcher, E. P. (1997). Cognitive biases in anxiety and attention to threat. *Trends in Cognitive Sciences, 1*, 340-345.

McDonnell, P. M. (1975). The development of visually guided reaching. *Perception & Psychophysics*, *18*, 181-185.

Miller, G. F. (2007). Sexual selection for moral virtues. *The Quarterly Review of Biology*, *82*, 97-125.

Miller, P., Eisenberg, N., Fabes, R. A., & Shell, R. (1996). Relations of moral reasoning and vicarious emotion to young children's prosocial behavior towards peers and adults. *Developmental Psychology, 32*, 210-219.

Moore, M. K., & Meltzoff, A. N. (1999). New findings on object permanence: A developmental difference between two types of occlusion. *British Journal of Developmental Psychology, 17*, 563-584.

Nesse, R. M. (2009). Social selection and the origins of culture. In M. Schaller, S. J. Heine, A. in Norenzayan, T. Yamagishi & T. Kameda (Eds.), *Evolution, culture, and the human mind*. Philadelphia, PA: Lawrence Erlbaum Associates.

Nesse, R. M. (1990). Evolutionary explanations of emotions. *Human Nature*, *1*, 261-289.

Nesse, R. M. (2007). Runaway social selection for displays of partner value and altruism.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-259.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs?. *Science*, *308*, 255-258.

Oster, H. (2003). Emotion in the infant's face. *Annals of the New York Academy of Sciences*, *1000*, 197-204.

Oster, H., & Ekman, P. (1978). Facial behavior in child development. In *Minnesota symposia on child psychology, 11*, 231-276.

Perkins, A. M., Inchley-Mort, S. L., Pickering, A. D., & Corr, P. J. (2012). A facial expression for anxiety. *Journal of Social and Personality Psychology, 102*, 910-924.

Perkins, A. M., Leonard, A. M., Weaver, K., Dalton, J. A., Mehta, M. A., Kumari, V., Williams, S.C. R., & Ettinger, U. (2012, October 1). A dose of ruthlessness: Interpersonal moral judgment is hardened by the anti-anxiety drug lorazepam. *Journal of Experimental Psychology: General*. Advance online publication. doi: 10.1037/a0030256

Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition, 36*, 1-6.

Premack, D., & Premack, A. (1997). Infants attribute value ± to the goal-directed actions of self-propelled objects. *Journal of Cognitive Neuroscience, 9,* 848–856.

Prinz, J. J. (2006). The emotional basis of moral judgments. *Philosophical Explorations, 9,* 29-

43.
Rozin, P., & Fallon, A. E. (1987). A perspective on disgust. *Psychological Review*, *94*, 23.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*, 145-172.

Sagi, A., & Hoffman, M. L. (1976). Empathic distress in the newborn. *Developmental Psychology*, *12*, 175.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, *300*, 1755-1758.

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin*, *34*, 1096-1109.

Schwarz, N. (1990). Feeling as information: Informational and motivational functions of affective states. Handbook of motivation and cognition.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*, 513.

Senju, A., Southgate, V., Snape, C., Leonard, M., & Csibra, G. (2011). Do 18-month-olds really attribute mental states to others? A critical test. *Psychological Science, 22*, 878-880.

Shweder, R. A. (1990). In defense of moral realism: Reply to Gabennesch. *Child Development, 61*, 2060-2067.

Simion, F., Regolin, L., & Bulf, Hermann. (2008). A predisposition for biological motion in the newborn baby. *PNAS, 105*, 809-813.

Spelke, E. S. (1990). Principles of object perception. *Cognitive Science, 14*, 29-56.

Sloane, S., Baillargeon, R., & Premack, D. (2012). Do infants have a sense of fairness? *Psychological Science, 23*, 196-204.

Storbeck, J., & Clore, G. L. (2008). The affective regulation of cognitive priming. *Emotion*, *8*, 208.

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, *54*, 768-777.

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie, 20*, 410-433.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46, 35-57.

Turiel, E. (1983). *The development of social knowledge: Morality and convention.*

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science, 311*, 1301-1303.

Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science*, *16*, 780-784.

Williams, G. C. (1966). 1996. *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton University Press.

Wilson, D. S., & Wilson, E. O. (2007). Rethinking the theoretical foundation of sociobiology. *The Quarterly Review of Biology*, *82*, 327-348.

Wolff, P. Observations on the early development of smiling. In: Foss BM, editor. Determinants of infant behavior. Vol. 2. Methuen; London: 1963. pp. 113–138.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*, 1-34.

Woodward, A. L.& Sommerville, J. A. (2000). Twelve-month-old infants interpret action in context. *Psychological Science, 11,* 73–77.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358*, 749-750.