How could someone who rewards me be wrong?

Young children's preferences for and attributions of competence to inaccurate rewarding versus inaccurate non-rewarding interaction partners

A dissertation by

Maria E. Barth

In partial fulfillment of the requirements for the degree of

The Doctor of Philosophy

in

Psychology

TUFTS UNIVERSITY

AUGUST 2015

ADVISER:

Nalini Ambady (Tufts and Stanford), in memoriam

Heather Urry (Tufts), interim adviser

Carol Dweck (Stanford), interim adviser

Abstract

Preschool-aged children are adept at detecting another person's obvious inaccuracy or ineffectiveness. Do the social emotional characteristics of an agent influence children's preferences for an agent? Do these characteristics influence children's assessments of an agent's competence? Previous research shows that adults who have high self-esteem are more likely than those who have low self-esteem to view individuals who provide favorable feedback as more competent; the current study examines whether this phenomenon is also present in children. The current study also examines whether, in addition to endorsing the competence of agents, children follow, or reject, the lead of an inaccurate agent in their behavior, a nonverbal measure of children's endorsement of an agent's competence. In two experiments, I examined children's perception of agents who act in a both socially-emotionally positive manner and also inaccurate manner. Experiment 1, a cross-sectional study, found that a higher age predicts choosing the positive inaccurate agent as a favorite, and that children of all ages are more likely to follow the incorrect lead of a positive inaccurate agent than follow the lead of a negative one. Experiment 2 found replication of the effect in Experiment 1 that higher age predicts choosing the positive inaccurate puppet as the favorite. Neither Experiment 1 nor Experiment 2 found evidence for self-esteem as a predictor of any of these outcomes.

Keywords: early childhood, person perception, trust, skepticism, persuasion, learning

Acknowledgements

Thank you to my parents and to my sister Stella for their support of my educational, professional, and personal pursuits over the many years since I decided, in high school, that I wanted to obtain a PhD in psychology. My family taught me that love and treating others with fairness, respect, and integrity are what are most important in life. My parents worked hard over many years so that I could have an easy life to pursue this level of education. Half my life has happened in the 15 years since I decided I wanted to pursue a PhD, and my family's love, care, encouragement, and warmth, and the worldview my family has taught me by example have guided me every step of the way. In addition to her constant support throughout my grad school program, thank you also to my sister Stella for helping create the puppets used in this research!

Thank you to my husband Eric for his support, from my undergrad honors thesis in psychology through my PhD program in psychology, and for sharing the date of our third wedding anniversary with my dissertation defense. His love, care, encouragement, warmth, and worldview have also been with me every step of the way. We've had many adventures, serious and silly, over the years, and I look forward to many more.

Thank you to the undergraduate research assistants at Stanford who worked on this project: Danielle Serrano, Jemima Oso, Nichole Kim, and additional assistance from Alberto Covarrubias. This project could not have happened without them.

Thank you to my dissertation committee, Prof. Heather Urry, Prof. Keith Maddox, Prof. Paul Muentener, and Prof. Carol Dweck, for their feedback, thought-provoking discussions, and, after the serious business was taken care of, sense of humor at my dissertation defense.

Thank you to Prof. Heather Urry of Tufts for taking me on as an interim advisee after Prof. Nalini Ambady passed away in 2013. Prof. Urry's advising and high standards have made me a better researcher and I am grateful, and will be for years to come, that I was able to complete my dissertation under her guidance. She set an example of what the research process can be like that I've been motivated to emulate. Thank you to the Emotion, Brain, and Behavior Lab (Urry Lab) for feedback on my dissertation talk and for the festivities at my defense.

Thank you to Prof. Carol Dweck of Stanford for allowing me into her lab as an interim student following Prof. Nalini Ambady's passing in 2013. Prof. Dweck has introduced me to research in education as well as psychology, and introduced me to many brilliant colleagues. I will remember research conversation over coffee with labmates fondly. The empowerment of a growth mindset for children, as well as for adults, is something I'll keep in mind, beyond the lab, in the future.

Prof. Nalini Ambady had the role of my adviser from 2010-2013, first at Tufts University, and then at Stanford University after she accepted a job at Stanford and we in the lab followed her to Stanford as non-matriculated students.

Thank you to the Interpersonal Perception and Communication Lab (Ambady Lab) at Tufts and Stanford for feedback on the early stages of these experiments, especially Dr. Brent Hughes who advised me on this project during Prof. Ambady's illness.

Thank you to the administrative, information technology, financial, maintenance, and custodial staffs who made conducting research possible at Tufts University and Stanford University. From Tufts 2010-2012: Cynthia Goodard, Lidia Bonaventura, and Madeline Amico (administrative), Jon Doorneweerd (information technology), Jusea and Ana (custodial). From Stanford 2012-2015: Roz Grayson, Maureen Sullivan, Porsche Williams, and Lynda Ichinaga (administrative), Harry Bahlman (maintenance), Ipshita Sengupta and Bhavna Raval (financial).

Thank you to the professors who previously supervised me as a research assistant or lab manager who inspired me to pursue graduate studies in psychology and shaped my research interests and trajectory: Prof. Elizabeth Spelke of Harvard on studies of infants' and young children's number cognition and social cognition, Prof. Susan Johnson of Stanford and Ohio State on studies of infants' attachment to their primary caregiver and expectations of symbolic representations of attachment as predicted by oxytocin transporter genotype, and Prof. Vinod Menon of Stanford School of Medicine on fMRI studies of mathematical and social reasoning by children who have mathematical learning disabilities and autism spectrum disorders.

Thank you to my own teachers: Basic Beginnings and Purpose School nursery schools, the Stoneham, Massachusetts Public Schools K-12, and the Tufts University Summer School program 2001. It was at the Tufts summer program, my first time on a college campus, where I paid part of the tuition from my summer job at Stop & Shop Supermarket and my parents paid part, that I first saw in practice that experimental psychology is something that people do for a living, and I thought I might want to pursue it someday.

Thank you to all of my instructors, advisers, mentors, and peers at Harvard University, where I embarked on this journey of studying psychology at the university level and had many intellectually formative, exciting, and memorable discussions with mentors and peers in labs, seminars, office hours, cafes, dorms, and dining halls. Several grad student and alumnae mentors in the Radcliffe Mentor Program and Women in Science at Harvard-Radcliffe also encouraged me on this path. I've sought to emulate them in my own mentoring roles for undergraduates at Tufts and Stanford. I am grateful to Harvard for its generous financial aid program that allowed my sister and I to work our ways through college, without placing a heavy financial burden on our parents or our future selves.

Thank you for the formal and informal education from all whom I've met within and outside the university whose demographics, backgrounds, outlooks, and experiences and mine differ. Every conversation has taught me that two people looking at the same thing can each see a different range of nuances and consequences in it that we can teach each other. We all bring our worldviews to our psychological research, and I think scientific objectivity requires exposing ourselves to varying viewpoints. The more worldviews represented in a room, the greater the range of human knowledge to draw from. There are voices from those who don't have access to the university that aren't heard in the university as much as they should be. We all stand to benefit greatly when we seek out our neighbors as equals, to listen and learn from them as we formulate our research questions.

Finally, thank you to Bing Nursery School, Stanford, CA, especially Jennifer Winters, Beth Wise, Chia-wa Yeh, all the head teachers (Peckie Peters, Parul Chandra, Mark Mabry, Adrienne Lomangino, Nancy Howe, Sarah Wright, and Todd Erickson), assistant teachers (too many to name but I would like to name them all!), staff, parents, and most of all, the children who participated in this research. It was a joy to interact with the children in their classrooms, get to know many of the parents, and interact with early childhood education professionals who create an environment where the children thrive. Prof. Eleanor Maccoby said that "Bing Nursery School is a national treasure" and I am grateful I was able to see it and be a part of it for a couple short, but intellectually and emotionally engaging, years. If we put our financial resources as a society into ensuring that every child, regardless of family income or race, had an education surrounded by adults who care deeply about them, interacting with friends from different walks of life, with ample indoor and outdoor space to explore, and materials that educate them about their world and inspire their imaginations, and every child encouraged to communicate his or her voice expecting that it will be heard and respected, we'd be in much better shape as a society.

Table of Contents

Abstract……i

Acknowledgements……iii

Experiment 1

Introduction……1

Method……10

Design, Analysis, and Power……20

Results……20

Discussion……29

Experiment 2

Introduction……34

Method……35

Design, Analysis, and Power……36

Results……36

General Discussion……39

References……49

Appendices……55

How could someone who rewards me be wrong?

Young children's preferences for and attributions of competence to inaccurate rewarding versus inaccurate non-rewarding interaction partners

The preschool years are a period of growth in person perception. Ages 3 through 5 are a period of both rapid cognitive maturation, and for many children, a period of new experiences of interacting with more adults, older children, and peers, in more varied settings than home or a small number of caregivers. Adults vary in their interaction style and enforcement of norms, and peers and older children vary in their interaction styles. A preschool-aged child gradually starts to experience more varied feedback in factual and academic domains, from a wider variety of peers and adults.

Children are active participants in constructing knowledge from this feedback. When faced with information in a new situation, children may engage both intrapersonal evaluation - how they feel they performed – as well as interpersonal evaluation – how they assess another's competence in providing that feedback. When presented with information that is ambiguous as to whether it is accurate or not, children may be especially likely to engage these processes to resolve the discrepancy between what they think is true and what an interaction partner appears to think is true. This type of metacognitive comparison of one's own and others' competence may involve children's positive or negative feelings toward their own competence in general, perhaps captured by the children's general feelings of self-esteem, as well as feelings of favorability toward the interaction partner's competence as well as general social demeanor.

To set the foundation for asking these questions about how children resolve discrepancies about how they feel they performed versus feedback about how they performed, the following review will examine preschool-aged children's sensitivity to others' inaccuracy in their person perception (Pasquini, Corrivequ, Koenig, & Harris 2007, Corriveau, Meints, & Harris 2009, MacDonald, Schug, Chase, & Barth

2013, Nurmsoo & Robinson 2009), preschool-aged children's tendency to be more affected by social-emotional information about a person more than competence information about a person (Brosseau-Liard & Birch 2010, Lane, Wellman, & Gelman 2013, Cain, Heyman, & Walker 1997, Stipek & Daniels 1990), preschool-aged children's tendencies to take positive feedback about their performance more into account than negative feedback (Stipek, Roberts, & Sanborn 1984, Bjorklund & Green 1992), and children's trait self-esteem in early childhood (Harter & Pike 1984, Davis-Kean & Sandler 2001). The literature above sets the stage for the current set of two experiments, which asks two questions: (1) Does the favorability of feedback influence whether the child prefers an agent and also sees an agent as competent? (2) Are children who have higher self-esteem more likely to favor an agent who provides favorable feedback, even inaccurately, and also endorse such an agent as competent?

Young children's competency detection

Regardless of the favorability or unfavorability of information, young children might expect that most of what adults tell them about the world is factually correct, though this is not always the case. Education and socialization require transmitting knowledge, especially through teaching (Kim, Kalish, Weisman, Johnson, & Shutts 2015). Children, who are at an epistemic disadvantage compared to the adults and older peers around them, generally stand to gain immensely from this information transmission (Koenig and Jaswal 2013). The child's dependence on adults and more knowledgeable peers for information does leave the child vulnerable to misinformation, however, and children appear to be cost-oriented in these decisions (Koening & Doebel 2013). Very young children are able to detect obvious incompetence, such as mislabeling common objects, and fail to learn from inaccurate agents (Pasquini, Corriveau, Koenig, & Harris, 2007, Corriveau, Meints, & Harris 2009, MacDonald, Schug, Chase & Barth 2013, Nurmsoo & Robinson 2009). Incompetence, when demonstrated in a domain that preschool-aged children can understand, is highly salient to children. A stronger weighting of

incompetence-suggesting information rather than competence-suggesting information may serve as a protective mechanism against misinformation (Koening & Doebel 2013).

Many studies suggest that young children of preschool age will reject information from a speaker after just a few instances of errors of pedagogical information, such as naming an object (Pasquini, Corriveau, Koenig, & Harris, 2007, Corriveau, Meints, & Harris 2009, MacDonald, Schug, Chase & Barth 2013, Nurmsoo & Robinson 2009). This avoidance of incompetent sources is further emphasized by children's failure to learn from incompetent sources on future lessons (Corriveau, Meints, & Harris 2009) and remembering a speaker's incompetence after a week has passed (Sabbach & Shaffman 2009), suggesting that children form a lasting impression of the speaker. Children as young as 3 distinguish inaccurate from accurate information, and children become more calibrated in distinguishing accurate information from irrelevant information as they get older (Corriveau, Meints, & Harris 2009).

Attribution of competence to another individual, depending on context, may require mentalizing abilities about the source of an agent's knowledge that develop later than the preschool years. There are no age differences in children ages 3-5 trusting a speaker who gives accurate names for objects while able to fully see them, over one who only makes mistakes while blindfolded suggesting that children might not make mentalistic interpretations of the speakers' inaccuracy at this age, possibly because pedagogical cues like eye contact and joint attention are absent (Nurmsoo & Robinson 2009). This finding may be an important caveat that if a task requires mentalizing abilities about the source of an agent's knowledge, children ages 3 through 5 may perform similarly on the task.

Generalizing competence to other personal qualities: developmental differences

While no previous studies have examined preschool-aged children's skepticism toward givers of inaccurate, yet favorable, feedback directly, several developmental studies have examined children's person perception of competence and other positive social characteristics separately.

There is evidence that young children may first develop the tendency to generalize positive social-emotional qualities to competence, and then around age 4 or 5 also begin to generalize the other way around, namely from competence to positive social-emotional qualities (Brosseau-Liard & Birch, 2010, Lane, Wellman, & Gelman 2013). One interpretation of this set of findings is that social-emotional traits may be more salient to younger children initially, and only later do they perceive competence as a desirable characteristic in its own right. A very young child's person perception may be more focused on whether the social-emotional characteristics of a person are favorable or unfavorable. This is in line with previous literature that infants', toddlers', and very young children's reasoning focuses on how someone makes them feel to gradually seeing people as psychologically complex by later in the elementary school years (Vaish, Grossman, & Woodward 2008). The current work investigates children's detection of competency and how children weigh competency versus the positive or negative valence of information the agent is providing. Positive or negative feedback given directly to the child is one way to do this.

In the above studies, either competence information or personal quality (e.g. prosocial, athletic) information was given. These studies did not combine multiple valences (e.g. incompetent but nice) in an individual to see which traits most affected the children. These traits were also descriptive and did not come across in direct interactions contingent on the child's actions. Experiments 1 and 2 vary both competence and whether the feedback provided is positive or negative toward the child.

Self-worth and assessing feedback-givers: Examples from the adult literature

One's self-schema impacts how we view others, particularly how we interpret ambiguous situations. Some examples from developmental literature include the "hostile attribution bias", where aggressive school-aged children are more likely to interpret a peer's ambiguous actions as aggressive (Dodge & Tamlin 1987), and the "internal working model of attachment" where securely-attached infants expect parents and children to be responsive to one another and insecure-avoidant infants expect parents and children to avoid one another, and respond with surprise when they act differently (Johnson, Dweck, & Chen 2007; Johnson, Dweck, Chen, Stern, Ok, & Barth 2010). Could a child's self-schema for perceived self-competence lead to a biased interpretation of inaccurate feedback he or she receives? For example, would a child who is high in perceived self-competence think he or she did well on a task after false positive feedback? Additionally, could a child's sense of self-competence moderate the interpretation of the competence of others, so that a child with a high perceived level of self-competence would think that those who provide false positive feedback are competent?

Turning, briefly, to the adult literature, many studies have shown that adults are not unbiased truth-seekers when it comes to interpreting feedback and self-esteem heightens these effects. One's sense of self-esteem might moderate the reaction to positive or negative feedback and the attributions one makes to the messenger depending on whether he or she likes the message. One's desired feelings toward himself or herself might be so strong as to lead to endorsing untrue statements about himself or herself. Indeed, in the general adult non-clinical population, adults tend to have an inflated positive view of the self (Taylor & Brown 1988,Sedikides & Gregg 2008). The average adult thinks he or she is an above-average driver (Svenson 1981), claim credit for successes while renouncing responsibility for failures (Mezulis, Abramson, Hyde, & Hankin 2004), and capable of controlling outcomes that are actually determined by luck (Langer 1975). When faced with congenial information relevant to themselves, adults tend to think "Can I believe it?" perhaps invoking trust, while for uncongenial information think, "Must I believe it?" perhaps invoking skepticism (Dawson, Gilovich, & Regan 2007).

Negative information may elicit more cognitive processing in general than positive information (Vaish, Grossman, & Woodward 2008). The typical adult is more likely to endorse positive than negative feedback about himself or herself from an outside observer (Brown 1986).

Adults who have high self-esteem are particularly more likely to show many of these effects (Sedikides & Gregg 2008), including those who have high self-esteem evaluating a feedback-provider as competent when the feedback was positive, and evaluating a feedback provider as incompetent when the feedback was negative (Kernis & Sun 1994). Thus, self-esteem may be a viable candidate among types of self-schemas to consider in how children might interpret inaccurate feedback and rate the competence of others.

Children, of course, are in a position to be more trusting of adult feedback, and it is generally advantageous for them to do so; certainly the status of the speaker and the context of the task matter. When the inaccuracy of the feedback is obvious, however, as it is in much inaccurate agent work with preschoolers, it is interesting to examine whether a preference for self-benefitting information, even if wrong, is present that may cause children to label a congenial yet inaccurate agent as competent.

Young children's self-worth and expectations of own performance

There are quantitative differences among young children, older children, and adults in their levels of self-esteem. Self-esteem fluctuates within individuals across the lifespan; self-esteem in early childhood tends to be high, declines during childhood, reaches a low point in adolescence, and gradually climbs again during adulthood (Robins & Trzesniewski, 2005). Young children, as a group, feel more positively toward themselves and their abilities, in an undifferentiated manner, than do adults. Several studies show that very young children often have an exaggerated sense of self-efficacy (Bjorkland & Green 1992). There is some evidence that preschool-aged children hold exaggerated expectations for their own performance (Stipek, Roberts, & Sanborn 1984). Four-year-olds had higher expectation of

themselves getting a reward than others getting a reward and children have higher expectations for their own performance after failure than others' performance after failure (Stipek, Roberts, & Sanborn 1984). It appears that children may focus on their desires for a reward or pleasant experience and not fully consider the skill required, and the cause and effect of the action of carrying it out (Stipek, Roberts, & Sanborn 1984). Children's ratings of their own abilities declined as they grew older, and children always rated themselves as better than peers regardless of age (Stipek & Tannat 1984). Children become more calibrated in their self-assessments across the elementary years.

However, given the overall exaggeratedly-positive sense of performance in children in this age group, high self-esteem is not universal. Children's self-concept is not as fully developed as adults', but some measures (Harter & Pike 1984) have demonstrated variation in self-esteem in preschool-aged children. These studies suggest that children of preschool age do not always separate their desires from their estimations of performance. Particularly when a reward is involved, children this age may not separate their desire for a reward with their real-world capabilities to bring the reward about. These findings lead to the question of whether, in connection with findings that children might endorse a person with positive social-emotional traits as competent even when the person doesn't actually have access to the information in question to be capable of competence (Lane, Wellman, & Gelman 2013), young children might not separate their desire for a reward with the circumstances produced by others to bring it about. Another question is whether children who have particularly high levels of self-esteem, like adults, are most likely to show these tendencies.

Summary

Preschool-aged children, as demonstrated in a number of studies (Pasquini, Corriveau, Koenig, & Harris, 2007, Corriveau, Meints, & Harris 2009, MacDonald, Schug, Chase & Barth 2013, Nurmsoo & Robinson 2009), are capable of detecting incompetence in others. However, a number of studies also

suggest that the youngest preschool-aged children may be more affected by the positive or negative social-emotional qualities, that is, behaviors that are communicative of feelings, of an agent than his or her cognitive competence and have a less-developed understanding of what competence entails (Brosseau-Liard & Birch 2010, Lane, Wellman, & Gelman 2013, Cain, Heyman, & Walker 1997, Stipek & Daniels 1990). No studies to date, however, have directly manipulated both in one agent. That is, no studies to date have manipulated both the positive or negative valence of information an agent provides as well as his or her competence, to see whether valence or accuracy more affects children's judgments of competence.

As a group, preschool-aged children tend to take information suggesting the ineffectiveness of their performance less into account in their self-assessments than do older children and adults (Stipek, Roberts, & Sanborn 1984, Bjorklund & Green 1992). Therefore, a child may be more affected by feedback in favor of his or her own performance than feedback suggesting he or she performed poorly. This effect might interact with age during the preschool years as well, such that children may show these effects more at certain ages.

There is variation in self-perceived competencies in young children (Harter & Pike 1984), but no studies to date have examined how a highly positive self-schema might moderate children's views of the competence of others. If young children are presented with the challenging situation of desirable feedback that is actually inaccurate, it remains to be seen how children will resolve the discrepancy: whether children will respond based on what they believe to be true, versus what the interaction partner appears to believe is true, given the interaction partners' errors are in their favor, and whether high self-esteem may moderate their responses.

Taken together, these findings could suggest that children, if like adults, might be more likely to rate as a favorite, endorse the competence of, and follow the lead of a positive inaccurate agent, and

question the competence of an unfavorable agent more than a favorable one. Children may ask "Can I believe it?" for positive, though inaccurate feedback, and ask "Must I believe it?" and expend more cognitive effort when feedback is negative (Dawson, Gilovich, & Regan 2007). These findings also suggest that high self-esteem could potentially function similarly in young children as in adults, in that high self-esteem might amplify these effects.

Experiments 1 and 2 aim to address these hypotheses: (1) The favorability of feedback influences which agent the child prefers and whether the child sees an agent as competent; specifically, children will be more likely to choose a positive but inaccurate agents as a favorite as compared to a negative and inaccurate agent and even a fully-accurate agent, and rate a positive but inaccurate agent as competent more than a negative but inaccurate agent, and possibly more than a fully-accurate agent. (2) Children with higher self-esteem may be more likely to choose an agent who provides favorable feedback, even inaccurately, as a favorite, and endorse an agent who provides favorable feedback, even inaccurately, as competent. (3) There are developmental differences in children's tendency to demonstrate (1) and (2). Specifically, given their less calibrated self-views and views of others, younger preschoolers may be more likely to demonstrate these tendencies than older preschoolers. Finally, (4) A longitudinal study might reveal whether children choose the same or a different agent as favorite or most competent at both time 1 and time 2, and whether children choose the same or a different agent at both time points depends on their age at time 1. Experiments 1 and 2 present preschool children with interaction partners who provide or deny rewards accurately or inaccurately to children while they play a game. Children's verbal and behavioral responses toward interaction partners are assessed. Experiments 1 and 2 also assess children's trait self-esteem using a standardized measure, and children's evaluations of their own performance on the task.

Experiment 1

The aim of Experiment 1 is to examine whether the valence of inaccurate feedback influences preschool-aged children's assessments of the feedback-giver's competence. Using a within-subjects design, participants, ages 3-5, played a card game with three puppets as informative agents, as many "inaccurate informant" studies with this age group use puppets as agents. Thus, there are three levels of feedback. One puppet always gave accurate feedback, rewarding children for a match on the card game and withholding reward for a mismatch, while the other two gave inaccurate feedback – one puppet always rewarded and one puppet never rewarded. Sessions were videotaped so reliability coders could assess the children's responses. Advantages of videotaped coding over "live" coding include a check against experimenter error if coding "in the moment", a check against any experimenter bias in coding, even if a non-deliberate bias, and a less interrupted and therefore more engaging game experience for the child. Following their interactions with all three puppets, children were asked to pick which puppet is most knowledgeable about the card game and which was their favorite. This measure provided information about young children's relative weighting of positive and negative self-relevant information, whether similar across ages 3-5, or showing a developmental change.

A second aim of Experiment 1 is to address the hypothesis that children who have higher trait self-esteem may be more likely to accept feedback that is in their favor, even if it is inaccurate. Children who completed the inaccurate agents task outlined above were also invited to complete a self-esteem measure (Harter & Pike 1984) in a separate session. Experiment 1 will provide information about whether differences in self-esteem influence children's preference for and perception of competence in agents who provide them with task feedback.

Method – Experiment 1

Participants

Children attended a mixed-age (3-5 years) classroom in a university-affiliated lab school with a play-based curriculum in the western US. In determining the sample size, the aim was to recruit 16-20 three-year-olds, 16-20 four-year-olds, and 16-20 five-year-olds, sample sizes similar to other studies testing age differences in children's assessments of competence (Kim, Kalish, Weisman, Johnson, & Shutts 2015, MacDonald, Schug, Chase & Barth 2013). In total, 102 children were recruited (61 females; 35 three-year-olds, 21 females; 44 four-year-olds, 27 females; and 23 five-year-olds, 13 females) for the inaccurate agent session; a smaller number (n=60; 41 females; 19 three-year-olds, 14 female; 27 four-year-olds, 19 female; and 14 five-year-olds; 8 females) participated in the follow-up self-esteem session. Recruiting was done in the classroom where children are allowed out of the room for 15 minutes at a time per day; thus the inaccurate agent and self-esteem sessions were completed separately. Parents gave prior consent for their children to participate and media release for videotaping, and children gave assent. Upon giving assent, children went to a separate testing room with the experimenter. The experimenters for Experiment 1 were all female: 1 White, 1 Black, 1 East Asian, and 1 Latina (for early pilot sessions) experimenters.

Materials and Procedure

Performance task. A memory card game, "Concentration", was chosen as the performance task in this study because it is familiar to many children this age, and also so that it would be obvious to the child whether the agent's feedback on whether shapes match is accurate or inaccurate. A game was chosen as opposed to an academic task, as an academic task may lead children to feel social pressure to endorse feedback as competent. While children in previous inaccurate agent studies labeled inaccurate agents as incompetent in pedagogical contexts, these studies did not involve feedback contingent on the child's actions. The interaction was "live" rather than pre-recorded to produce ecological validity for the child.

Familiarization trial. To familiarize children with the memory card game, the experimenter presented the child with four cards showing two matching pairs of shapes and asked the child to identify which shapes match. The experimenter instructed the child that in this game, matching shapes are called "a match" and are put in a "match pile" to the child's right-hand corner of the table. Once children reliably demonstrated mastery of following these instructions, the experimenter presented the child with face-down cards and instructed the child to turn two over at a time to find a match. Upon finding a match, children were instructed to put the cards in the match pile. Upon turning over two cards that are not a match, children were instructed to put the cards back face-down and try again. Once the child demonstrated mastery of following these instructions, the experiment began. This action of putting matching cards in a match pile or non-matching cards back in the non-matches piles is assessed later in the experimental session as a dependent variable to determine if children were more likely, behaviorally and nonverbally, to follow the inaccurate lead of a positively-valenced agent by putting the cards in the wrong pile, where the puppet inaccurately thinks they belong, rather than where the child thought they belong, than for the lead of a negatively-valenced agent.

Agent introduction. The experimenter then told the child that he/she would play the exact same card game he or she just played with the experimenter, but this time with three puppets. The experimenter set up a foam puppet stage and three felt puppets, red, green, and blue [Figure 1]. The experimenter said the puppets all "say and do different things", and asked the child if he or she had a favorite book or movie and asked the child for the names of characters in the book or movie. If a child did not identify a favorite book or movie, the experimenter chose *Winnie the Pooh* or other well-known children's story, or a book she knew was currently being read regularly in the child's classroom. The experimenter said that the characters in the child's favorite book or movie all say and do different things, and the puppets are like this. The experimenter then removed the green and blue puppets and said, "Let's see what this puppet has to say." Puppets were always presented in the order red-green-

blue, with positive inaccurate and negative inaccurate behavior counterbalanced between green and blue, in the assumption and hope that children with a favorite color preference would be randomly distributed across those two conditions. The order red-green-blue was presented for all children with the expectation that sample sizes would be small and so it would be difficult to test for order effects with few children in each condition. The accurate puppet was always presented first, so that it would be seen as precedent-setting and a baseline of the correct way to play the game, and the inaccurate puppets were always presented following the accurate puppet in contrast to it [Figure 2]. The aim was to have children view the inaccurate puppets as inaccurate and not simply as playing the game by different rules.

Accurate "baseline" agent. The experimenter voiced the red puppet in an engaging and encouraging tone, "Hi there. We're going to play the card game. Every time you find a match, I'll give you a teddy bear! You can put the teddy bears in this teddy bear house! Go ahead and try to find a match!" and presented the child with a toy house made out of a strawberry container and construction paper. The experimenter then gave the child 6 face-down red cards arranged in a 3x2 pattern. The child then completed the memory card game. Each time the child finds a match, the red puppet said, "Yay! It's a match! Here's a teddy bear for you!" and presented the child with a small red plastic teddy bear (commonly used in pre-K and kindergarten classrooms to learn counting and arithmetic, but not used at the child's school) velcroed to the puppet's hand, which the child was then encouraged to take off and put in the teddy bear house. Each time the child does not find a match, the red puppet says, "No, not a match. No teddy bear. Try again!" The experimenter provided any guidance in case the child did not follow the task accurately (e.g. attempted to turn over more than two cards at once), but otherwise did not give the child any hints or guidance to complete the task. The experimenter did remind the child, if he or she did not put matches in the match pile, or returned non-matches to the non-match pile, to do so.

Comprehension check. Once the child has found all three matches from the card game and put them in the matches pile, the experimenter said, "Now I have some questions about the red puppet!" The experimenter asked two comprehension check questions: (1) "When you had two cards like this [presented two matching cards to the child], what did the red puppet do?" [Examples of accurate answers were "Gave me a teddy bear" or "He said, 'Yay it's a match".] (2) "When you had two cards like this [presented two different cards to the child], what did the red puppet do?" [Examples of accurate answers were "Didn't give me a teddy bear" or "He said, "No, that's not a match."] If the child does not answer a question, the experimenter prompted, "Did the red puppet give you anything when the cards were like this?"

Counterbalance order. The green puppet was presented next, in counterbalanced order, so that it was either a "negative" or a "positive" puppet. The experimenter introduced the second [and third] puppets by saying, "Now let's see what the green [blue] puppet has to say. Remember, all the puppets say and do different things. I wonder if this puppet might do something different. Let's see." The green puppet introduced itself by saying, "Now we're going to play the card game. Every time you get a match, I'll give you a teddy bear. Here's my cards!" The green puppet had green cards and gave out green teddy bears. The blue puppet had blue cards and gave out blue teddy bears.

Negative inaccurate agent [2nd or 3rd agent, counterbalanced]. The negative inaccurate puppet thinks no pairs of cards are matches. Each time the child turns over cards, matching or non-matching, the puppet says, "No, not a match. No teddy bear. Try again!" Upon the puppet's first inaccurate instance, that is, the first time the child turns over a matching pair but the puppet said, "no, not a match" the experimenter asked the child, "Is that really, truly a match? Where do the cards go?" so that a measurement could be made of the children's following the incorrect lead of the puppet versus placing the cards where the child knew the cards should go. If the child did not put the cards in the

matching pile, this would count as following the puppet's inaccurate lead, the experimenter emphasized that the cards are really, truly a match, and instructed the child to put the cards in the matching pile. If the child protested or questioned the puppet's denial of a teddy bear, the puppet said, "Hmmm. No, I don't think those are a match. Those do not look the same to me."

Positive inaccurate agent [2nd or 3rd agent, counterbalanced]. The positive inaccurate puppet thought all pairs of cards were matches. Each time the child turned over cards, matching or non-matching, the puppet said, "Yay! It's a match! Here's a teddy bear for you" and gave the child a teddy bear. Upon the puppet's first inaccurate instance, that is, the first time the child turned over a non-matching pair but the puppet said, "That's a match!" and gave the child a teddy bear, the experimenter asked the child, "Is that really, truly a match? Where do the cards go?" so that a measurement could be made of the children's following the incorrect lead of the puppet versus placing the cards where the child knew they should go. If the child did put non-matching cards in the matching pile, this counted as following the puppets' inaccurate lead, the experimenter emphasized that the cards were not really, truly a match, and instructed the child to put the cards back in the active game pile. If the child protested or questioned the puppet giving a teddy bear, the puppet said, "Hmmm. I think those are a match. Those cards look the same to me."

Once the child found all three matches from the card game and put them in the matches pile, the experimenter said, "Now I have some questions about the green [blue] puppet!" The experimenter asked two comprehension check questions, matched pair and non-matched pair, as with the accurate baseline puppet above.

Child preference without experimenter demand. The experimenter then pretended that she forgot something in the cabinet, laid out all three puppets in front of the child, and said he or she could choose one puppet to play with while the experimenter was looking in the cabinet with her back turned

to the child. The experimenter returned to the table about 30 seconds later. The purpose of this measure was to see which puppet children may play with without experimenter demand; that is, children may think the experimenter wants the child to choose a particular puppet as his or her favorite, but this measure allows the child to play with a puppet without the attention of the experimenter.

Preference measure. The experimenter then said to the child, "Now I have some questions about these three puppets! Which puppet played the card and bear game the way you like the most?" If the child chose more than one puppet, the experimenter prompted the child to choose one. If the child insisted on choosing two after several prompts to choose one, the experimenter allowed it. Later data analysis separately grouped children who only chose one puppet from ones who chose more than one and excluded them from some analyses. The experimenter asked the child, "Why?" for the chosen favorite, and then took the puppet away, leaving the remaining two (or, in some cases, one). The experimenter then said, "So now we have these two puppets. Out of these two puppets, which puppet plays the card and bear game you like the most?" For the remaining puppet, the experimenter said, "Do you like how this puppet played the card and bear game? Why?" For any of the why questions, if the child gave justification that does not relate to the game, such as "He's my favorite color," the experimenter put the puppets back in a row and said, gently and enthusiastically, "Ah, I see. That's your favorite color. I have a *different* question! I want to know, which puppet *plays cards and bears* the way you like the most? Why?" until the child justified his or her answer with a reason beside a favorite color or failed to justify an answer repeatedly so that the experimenter made the decision to move on.

Competence measure. In an identical procedure, the experimenter then put all three puppets in a row and asked, "Which puppet *knows the most* about how to play the cards and bears game? Why?" The same procedures were followed in removing the chosen most competent and asking about the competence of the remaining two. Similarly, the child was re-asked the question "Which puppet

knows the most?" if he or she identified the puppet as knowing the most because it's his/her favorite color.

Self-assessment of performance measure. The experimenter laid all the puppets out in a row and asked the child if he or she was good at all the puppets' games, or not so good. The experimenter asked the child if he or she is better at some than others, asking "why" after the child's answers. The purpose of this measure was to determine whether children's view of their own performance was impacted by the agent's pattern of rewarding.

Self-assessment of difficulty of task measure. Similarly, the experimenter laid out all the puppets in a row and asked the child if any of the puppets' games were easier or harder than others, or if they were all the same, asking "why" after the child's answers. Similarly, the purpose of this measure was to determine whether children's view of the difficulty of the task was impacted by the agent's pattern of rewarding.

Final comprehension check. The experimenter took each puppet, one by one, and repeated the comprehension check with the child for a matched and non-matched pair per puppet, two pairs by three puppets equals six questions total. See Figure 3 for the full experimental design.

Self-esteem measure. Per school policy, children are allowed outside the classroom for research no longer than 15 minutes at a time. Therefore, all except one child completed the self-esteem measure in a separate session. All efforts was made to administer the session within one month to the child, but this was not always possible due to factors including child refusal to return to the research room a second time, school vacations, research assistant illness, and children's absences from school. Median number of days between puppets and self-esteem session = 14.00, modal number of days=5.00, mean number of days=27.43, SD=39.81.

The Harter and Pike (1984) "Pictorial Scale of Perceived Competence and Social Acceptance, Preschool-Kindergarten" measure was used, including all four subscales – cognitive competence, physical competence, maternal acceptance, and peer acceptance. The Cronbach's alpha for internal consistency of the scale is reported as "in the mid to high .80s" (Harter & Pike 1984, p. 1974). The scale has predictive validity in that children rated in the top or bottom quartile of their class by their teachers significantly differed in their cognitive competence scores on this scale (Harter & Pike 1984). It has shown discriminant validity in that children held back a year in school had lower scores on the cognitive subscale, children who had been born pre-term had lower scores on the physical subscale, children who scored higher on a measure of child depressive symptoms also scored lower on the maternal acceptance subscale (Harter & Wright 1984), and children who had just moved to a new school within 2 months had lower scores on the peer acceptance subscale (Harter & Pike 1984). The cognitive subscale had convergent validity in that 96% of children gave specific reasons to explain their self-ratings of cognitive competency (Chao, Harter, Adams, & Strop 1983). The reliability of this measure increased as children grew older, and whether this is due to developmental limitations or language limitations is unknown (Harter & Pike 1984, Davis-Kean & Sandler 2001). This measure has higher reliability with high-income children, as one study showed low-income children did not comprehend the Likert-like scale used in this measure (Fantuzzo, McDermott, Manz, Hampton, & Burdick 1996).

In the Harter and Pike (1984) measure, for each item, the child is presented with two black-and-white pictures of gender-matched children, one who is described as doing well at a task or feeling accepted, and the other described at not doing well at a task or not feeling accepted. For example, "This boy knows the first letter of his name" and "This boy has trouble remembering the first letter of his name" followed by "Which boy is most like you?" Once the child chooses, the experimenter follows up, pointing to a large circle and a smaller circle underneath the picture the child chose, "Are you really good or pretty good at remembering the first letter of your name?" and the child then points to the

circle indicating the degree of competence. All four subscales (peer, maternal, physical, cognitive) were administered, six questions for each subscale except for omitting the cognitive competence question about children receiving stars on their papers, as this method is not used at their school and would be unfamiliar to them, so that the cognitive measure is rated out of 5 items instead of 6. Therefore the range of possible scores are 23 (instead of 24) to 92 (instead of 96). Examples of cognitive competence were knowing the first letter of his or her name, knowing how to count, and knowing the names of colors. Examples of physical competence were being able to swing by him or herself on a swing, running, skipping, and climbing on monkey bars. For the competence questions, children chose a response from "Not at all", "Sort of", "Pretty good", or "Really good" to describe their abilities. Examples of maternal acceptance were mother smiles at child, mother talks with child, and mother reads with child. Examples of peer acceptance were peers invite child to play, child has many friends to play games with, and child has friends to play with on playground. For the acceptance questions, children chose a response from "Never"/"Hardly ever", "Sometimes," "Pretty much," or "A whole lot".

The Harter and Pike (1984) measure, when it was normed, revealed variation in preschool-aged children's self-esteem scores. Preschool-aged children were not universally at ceiling on this measure. Preschool-aged children showed, on the cognitive subscale, a mean of 3.4 out of 4 and standard deviation of 0.45; on the physical subscale, a mean of 3.2 and standard deviation of 0.49; on the peer subscale, a mean of 3.0 and standard deviation of 0.56; and on the maternal subscale, a mean of 3.1 and standard deviation of 0.59.

Inter-rater reliability. All experimental sessions were videotaped, and trained research assistant coders viewed the sessions in the lab to record the child's responses. Interrater reliability Cohen's Kappa overall was 0.90, with a kappa of 0.90 for the puppets task and 0.88 for the self-esteem task.

Design, Analysis, and Power

It is common for studies in the behavioral sciences to have small effect sizes (Cohen 1992, Faul, Erdfelder, Lang, & Buchner 2007) and all calculations were done with the power calculator by Faul, Erdfelder, Lang, & Buchner (2007) "G*Power" version 3.1.9. or IBM SPSS Statistics 22, Sample Power 3.0.1. This study, possibly the first to manipulate valence and accuracy separately in inaccurate informants, follows other developmental studies of inaccurate informants with sample sizes of around 20 per age group, and small effect sizes. It is typical in developmental studies, as a difficult to recruit population, to be underpowered. The results of this preliminary study may be used to design future studies with greater anticipated power.

The analyses in Experiment 1 are mostly logistic regressions on dichotomous criterion variables, whether the child chose or did not choose a particular agent to answer the given question. The predictor variables include age as a continuous predictor, total score on the Harter and Pike (1984) self-esteem scale, and interactions of these variables. These analyses report the log odds (standardized B) as well as the odds ratio (Exp(B)). All tests are two-tailed with an alpha of 0.05. Effect size and power for these logistic regressions are generally low, effect sizes below 0.3 and power below 0.2. A few analyses in Experiment 1 require t-tests, with effect sizes in the 0.2 to 0.3 range and power in the 0.50 to 0.60 range.

Results – Experiment 1

Influence of favorability of feedback on whether child sees an informant as likeable and competent.

Favorite agent. The hypothesis was that younger children may be more likely than older children to choose the always-rewarding, though inaccurate, agent as their favorite. The purpose of this outcome measure was to assess which agent the child prefers. Three logistic regressions were run for the

accurate, inaccurate positive, and negative inaccurate puppets with dichotomous criterion variables; 1 if the child chose the puppet and 0 if the child did not choose the puppet; age was included as a continuous predictor. Analyses were performed on the full sample of 102 children. Children could choose more than one puppet; thus, a child could have a score of "1" for more than one criterion variable.

The omnibus model for the positive inaccurate agent was significant, $\chi^2(1)=5.117$, $p<0.05$ explaining 6.9% of the variance (Nagelkerke R^2), that contrary to the hypothesis, higher age predicted the likehood to choose the positive inaccurate puppet as the favorite, $\beta=0.700$, $Exp(B)=2.013$, $p < 0.05$. Age was not a significant predictor of choice of accurate puppet nor of negative inaccurate puppet as favorite. Additionally, age was not a significant predictor of the child stating that all puppets were equally their favorite, or stating that any combination of two puppets were their favorite: accurate and positive, positive and negative, or accurate and negative.

Overall, in the group of 102 children, the positive inaccurate puppet was slightly preferred as overall favorite; please see Figure 4 for overall group means and Figure 5 for favorite by age cohort means. Children were encouraged to choose only one puppet to answer each question, but children were allowed to choose more than one, and indeed, 26 of the 102 children did choose more than one puppet for the "favorite" analyses. These children were included in the above analyses because this inclusion captured equivalences children may have seen among the agents that would not have been revealed by forcing a single-item response. To explore if the same significant finding of preference for positive inaccurate agent predicted by higher age occurred in the sample of 76 children who chose just one puppet, so all participants only had a score of 1 for one of the criterion variables, another three logistic regressions (accurate, positive inaccurate, and negative inaccurate) were run with age as a continuous predictor. In the group of 76 children who only chose one agent as favorite, the positive

inaccurate puppet was still the slight favorite. The logistic regression results showed a marginal rather than significant result that higher age predicted choosing the positive inaccurate agent, $\chi^2(1)=3.390$, p=0.066, explaining 5.9% of the variance (Nagelkerke R^2), β=0.625, Exp(B)=1.868, *p*=0.073.

No other analyses were significant or marginal: choosing accurate only, choosing negative inaccurate only, choosing both accurate and negative, both accurate and positive, both positive and negative, or choosing all three agents as favorite. Please see Figure 6.

Most competent agent. The hypothesis was that younger children may be more likely than older children to choose the always-rewarding, though inaccurate, agent as the most competent. The purpose of this outcome measure was to assess which agent the child thinks is most competent. Three logistic regressions were run for the accurate, inaccurate positive, and negative inaccurate puppets with dichotomous criterion variables; 1 if the child chose the puppet and 0 if the child did not choose the puppet; age was included as a continuous predictor. Analyses were performed on the full sample of 102 children. Children could choose more than one puppet; thus, a child could have a score of "1" for more than one criterion variable.

No analyses were significant. There was one analysis that showed a trend; the omnibus model for the negative inaccurate agent, $\chi^2(1)=3.195$, p=0.074, explaining 6.0% of the variance (Nagelkerke R^2), that lower age predicts choosing the negative inaccurate puppet as most competent, β=-0.800, Exp(B)=0.490, *p* =0.086. Children of all ages picked the accurate puppet as most competent in higher numbers than the other puppets.

Overall, in the group of 102 children, the positive inaccurate puppet was slightly preferred as overall favorite; please see Figure 7 for overall group means and Figure 8 for means by age cohort. Children were encouraged to choose only one puppet per question, but children were allowed to choose more than one, and indeed, 26 of the 102 children did choose more than one puppet for the "favorite"

analyses. These children were included in the above analyses because it captured equivalences children may have seen among the agents that would not have been revealed by forcing a single-item response. To explore if the same finding of preference for positive inaccurate agent predicted by higher age occurred in the sample of 76 children who chose just one puppet, so all participants only had a score of 1 for one of the criterion variables. In the group of 78 children, the majority chose accurate as the most competent puppet. The results showed, for age as a continuous predictor of choosing the negative inaccurate puppet as the most competent was also marginal, $\chi^2(1)=3.266$, $p=0.071$, explaining 7.1% of the variance (Nagelkerke R^2), that lower age predicts choosing the negative inaccurate puppet as most competent, $\beta=-0.816$, $Exp(B)=0.442$, $p=0.085$. Please see Figure 9.

Following lead of inaccurate agent. The hypothesis was that children may demonstrate their belief in the competence or trustworthiness of an agent by their nonverbal behavior of playing the game how the inaccurate agent suggested without questioning; in this analysis, the hypothesis was that children would place the cards where the positive inaccurate agent thinks they should go significantly more often than placing the cards where the negative inaccurate agent thinks they should go. In other words, the hypothesis was that children, as a group, would follow the lead of the positive inaccurate agent but not the negative inaccurate agent. As younger children might be more sensitive to the valence of the agent's feedback, younger children might be more likely than older children to follow the lead of the positive agent than the negative agent. At the inaccurate agents' first inaccurate statement to the child, video coders rated whether the child placed the cards where they actually belong, or where the inaccurate agent suggested they belong. A paired-samples t-test revealed that children were indeed more likely to place the cards inaccurately after the first incorrect reward from the positive inaccurate puppet than the first incorrect denial of reward from the negative inaccurate puppet, $t(101)=-2.098$, $p<0.05$. Please see Figure 10. Further logistic regression analysis to examine whether the child's age predicted the likelihood to follow the lead of the positive inaccurate agent was not significant.

Self-esteem. Children's mean self-esteem scores and standard deviations in Experiment 1 were comparable to the values found in Harter and Pike (1984) when the scale was originally normed, suggesting that there was sufficient variation in this sample to detect effects. Please see Table 1 with descriptive statistics for the self-esteem measure in Experiment 1. The hypothesis was that self-esteem would moderate the effect of valence manipulation on children's choice of competent puppet and favorite puppet; specifically, the hypothesis is that children with high self-esteem would choose the positive inaccurate puppet as most competent and most preferred agent. There was no main effect of children's total summary score on the Harter and Pike self-esteem scale predicting children's puppet choice for most competent or favorite agent. Choices of favorite or most competent agent also were not predicted by the interaction of self-esteem and age. Logistic regression analyses to examine whether the child's age or self-esteem predicts the likelihood to follow the lead of the positive inaccurate agent was also not significant. There was no main effect of children's total summary score on the Harter and Pike self-esteem scale predicting whether children followed the lead of the positive inaccurate nor the lead of the negative inaccurate agent. The interaction of self-esteem and age likewise did not predict following the lead of the positive inaccurate or negative inaccurate agent.

Other analyses.

Child's self-assessment. The hypothesis was that children may demonstrate their view of an agent as competent to provide feedback by stating they actually did better in the positive, always-rewarding puppet's game, regardless of how they actually performed on the game. The purpose of this outcome measure is to assess whether children assess their own performance contingent on the rewards given by the agent. Children were asked how well they thought they performed on each puppet's game. As children tend to have a high sense of self-efficacy, it is possible they would rate themselves highly on agents' games across the board and so a null result is possible.

The omnibus model for the positive inaccurate agent was significant, $\chi^2(1)=5.370$, at $p < 0.05$, explaining 22.0% of the variance (Nagelkerke R^2), and a non-significant trend for higher age to predict choosing the positive inaccurate agent's game as the one the child was best at, β=2.499, Exp(B)=12.174, p =0.073. Analyses with the accurate puppet, negative inaccurate puppet, and combinations of all puppets, accurate and positive, positive and negative, and accurate and negative were not significant.

Self assessment of difficulty. The hypothesis was that younger children would think the positive inaccurate puppet's game was easiest compared to older children. The purpose of this outcome measure is to determine if children's sense of the difficulty of the task is contingent on the rewards given by the agent in that game. Logistic regressions (for choosing the accurate, positive inaccurate, and negative inaccurate agent) revealed that higher age predicted stating that the positive inaccurate puppet's game was the easiest, β=1.503, Exp(B)=4.495, $p < 0.05$, omnibus $\chi^2(1)=7.460$, $p<0.01$, explaining 15.7% of the variance (Nagelkerke R^2).

Performance measure. The hypothesis was that children may take more card-pair turns to complete the task with inaccurate agents than accurate agents. Furthermore, the hypothesis was that children might take a different number of card-pair turns based on the valence of the inaccurate agents. The purpose of this outcome measure was to determine if children perform the game equally well with each of the three agents, as measured by the total number of card-pair turns and the ratio of errors to accurate pairs the child takes to complete the game. The number of card-pair turns was an approximation of the child's performance on the task. As both luck and skill are involved in the number of card turns it takes for a child to complete the task, it is not a perfect approximation. However, averaged across each puppet type in aggregate for all participants, this measure could be informative of any differences in children's performance across agents.

Each agent's card game involved six cards; i.e. three pairs. Though luck is involved in whether the first card-pair turn is a match or non-match, skill is involved to complete the task, as the child must remember where matching shapes are. Therefore, children took at minimum three card-pair turns to complete the task. When this happened, the experimenter had the child turn over an additional, fourth, pair of non-matching cards to see how the puppet reacts to the non-match. So any child, on any given agent's game, took at least four card-pair turns to complete the task. Children were allowed to continue the game for as many card-pair turns as needed to complete the game.

Children's performance on the task was measured in two different ways. The first was the raw number of card-pair turns the child took to complete the task with each puppet: at least four. The second performance measure is the ratio of non-matches (errors) to matches (hits) the child produced with each puppet. I analyzed the hypothesis of whether children take more card-pair turns to complete the task and had a higher ratio of errors to total turns with a linear mixed effects model ANOVA, using the accurate puppet as a comparison base and the outcome as error ratio as well as an outcome of number of card turns. There were no significant differences of error ratio among the different types of puppets. There was, however, a significant difference in total number of card-pair turns between the positive and accurate puppets, $t(101)=-2.889$, $p<0.01$. The average number of card-pair turns for the accurate puppet was 5.44, for positive inaccurate was 6.37, and for negative inaccurate was 5.95.

I ran a repeated measures ANOVA with effect coding to determine whether the mere presence of a reward impacted children's performance (accurate agent=1, positive inaccurate agent=1, negative inaccurate agent=-2), whether a linear effect of reward is present (accurate agent=0, positive inaccurate agent=1, negative inaccurate agent=-1), and whether the inaccuracy of the agent's statements impacts children's performance rather than the presence or absence of reward, e.g. distracted attention

regardless of valence (accurate agent=-2, positive inaccurate agent=1, negative inaccurate agent=1). None of these contrasts were significant.

This analysis also tested for order effects. It is possible that children became fatigued by the game by the time they played with the third puppet. In that case, the counterbalance order in which an agent appears would impact children's performance. Children had a higher error ratio for the third puppet they played with, compared to the second puppet, regardless of puppet valence, $t(202)=2.244$, $p<0.05$. This effect seems to be driven by the positive inaccurate puppet; children did not perform significantly differently for the negative inaccurate puppet regardless of whether the child saw the negative inaccurate puppet second or third; however, children had a higher error ratio with the positive inaccurate puppet when the child played with the positive inaccurate puppet third instead of second, $t(100)=-1.995$, $p<0.05$.

Puppet choice without experimenter demand. The hypothesis is that children will choose to play with the puppet they prefer, and this method may capture children's favorite puppet more than the question "Which puppet is your favorite?" above, as the experimenter's back is turned so there is no experimenter demand. The purpose of this outcome measure was to assess, through children's nonverbal behavior, immediately after playing with all three puppets, which puppet children choose to play with in the absence of social demand of the experimenter's questioning. Children's choice of puppet to play with while the experimenter's back is turned was coded for both the child's first choice and then final choice, if the child played with more than one puppet. Logistic regressions examining age as a predictor, self-esteem as a predictor, and the interaction of age and self-esteem as a predictor of puppet choice were all produced non-significant results.

Comprehension check 1 and Comprehension check 2. The comprehension checks were used to determine whether the children understood each puppet's pattern of rewarding or not rewarding

behavior. Comprehension check 1 occurred immediately after the child played with each puppet, the experimenter showed the child the puppet, a matching pair of cards, then a non-matching pair of cards, and asked what the puppet did for each case. The child's response was recorded. The child was corrected if he or she did not correctly identify what the puppet did in each case. Comprehension check 2 was an identical procedure, asked at the very end of the experimental session as the last questions asked of the child before leaving the research room and returning to his or her classroom.

Therefore there were 12 comprehension check questions: memory for agent type (3 levels; accurate, positive, negative), comprehension check type (match, no match), comprehension check timing (immediately after playing with puppet or at very end of game). I ran a logistic regression on each of the 12 questions as a criterion variable, with age as a continuous predictor. 3 of the 12 analyses were significant and 2 showed a trend; in all significant and trending cases, higher age predicted describing the puppet's behavior as it actually occurred.

The significant outcomes were that higher age predicted stating immediately after playing with the accurate puppet that it did not give a reward for non-matches, $\chi^2(1)=4.681$, $p<0.05$, explaining 9.5% of the variance (Nagelkerke R^2), $\beta=1.082$, $Exp(B)=2.950$, $p < 0.05$, higher age predicted stating immediately after playing with the negative inaccurate puppet that it did not give a reward for matches, $\chi^2(1)=14.374$, $p<0.01$, explaining 33.5% of the variance (Nagelkerke R^2), $\beta=3.043$, $Exp(B)=20.960$, $p < 0.01$, as well as stating the same for the negative inaccurate puppet in the very final question at the end of the game, $\chi^2(1)=8.914$, $p<0.01$, explaining 15.5% of the variance (Nagelkerke R^2), $\beta=1.090$, $Exp(B)=2.976$, $p < 0.01$.

The non-significant, but trending, outcomes, were that higher age predicts stating that the positive inaccurate puppet rewarded non-matches, $\chi^2(1)=3.790$, $p=0.052$, explaining 5.6% of the variance (Nagelkerke R^2), $\beta=0.650$, $Exp(B)=1.916$, $p =0.058$, and higher age predicted stating the same

for the positive inaccurate puppet when asked at the very end of the game, $\chi^2(1)=3.514$, p=0.061, explaining 6.0% of the variance (Nagelkerke R^2), β=0.620, Exp(B)=1.858, *p* =0.067.

Discussion

The goal of Experiment 1 was to assess whether preschool-aged children's preferences for and judgments of others' competence were affected by the valence of inaccurate information and whether these judgments track with age and self-esteem differences. Specifically, Experiment 1 aimed to assess whether children were more likely to prefer to play with, and endorse as competent, an interaction partner who provided positively-valenced inaccurate information than an interaction partner who provided equally inaccurate, but negatively-valenced information.

Previous research has demonstrated that children of preschool age assess agents as competent or incompetent based on pedagogical information they provide. To date, no known studies have separately manipulated the valence and accuracy of information provided to the children. The current work builds on previous findings of children's accuracy detection to assess whether children are less sensitive to inaccuracy when the information is in their favor, and whether there are any developmental differences in sensitivity to inaccurate positive versus negative information. There was some evidence for developmental differences in sensitivity to inaccurate positive versus negative information; there was a positive relationship with child's age predicting choosing the positive inaccurate puppet as favorite. There was a lack of support for the hypothesis that high self-esteem would predict children's choice of favorite or most competent agent.

There was a cross-sectional age difference in sensitivity to the valence of the agents' feedback, though it was in the opposite direction of the original hypothesis. Specifically, there was a tendency for older age to be associated with choosing the positive inaccurate agent as the favorite, though not as the most competent. Additionally, higher age predicted self-report as being "best at" the positive

inaccurate puppet's game compared to the other agents' games. It is possible that older children may have more experience with feedback contingent on their actions, and therefore especially value rewards more than younger children, based on their everyday experiences. The descriptive statistics of favorite puppet choice are more differentiated among the three puppets for the older children, with positive as the favorite, while more evenly divided among the younger children. It is possible that the older children are differentiating among the different puppets more readily than the younger children, and valuing rewards more than the younger children.

There was a lack of support for the hypothesis that children with high self-esteem would be more likely to choose the positive inaccurate agent as the most competent or their favorite. Self-esteem did not predict children's choice of favorite or most competent puppet. There may be several reasons for this. First, other constructs related to children's self-worth, rather than self-esteem, may be more relevant to children's choice of most competent or favorite agent. Though it is not directly analogous to self-esteem, at least as measured in this study, previous studies by Burhans and Dweck (1995) showed that while young children as a group tend to, initially at least, rate their performance on tasks highly, young children with a contingent sense of self-worth will, following negative feedback, downgrade those initially high views of their own performance on a task to take on a low opinion of their own work, while children whose sense of self-worth is not contingent will maintain their high opinion of their own work, even after receiving negative feedback on it. A future study could measure how children with contingent vs. non-contingent self-worth view their performance on a task after receiving positive or negative inaccurate feedback.

Another possible reason why self-esteem did not significantly predict puppet choice is that trait self-esteem simply might not relate to how children interpret inaccurate feedback in this task. Trait self-

esteem, measured on a different day from the agents task, might not be as sensitive a measure as state self-esteem measured on the same day as the agents task, for children this age.

One behavioral measure of whether the children questioned the competence of the inaccurate agents is, upon the inaccurate puppet making its first error, whether the children placed their cards where the cards actually belonged, or where the puppet incorrectly suggested the cards should go. Children followed the lead of the positive inaccurate agent more than the negative inaccurate agent, suggesting that children question the negative inaccurate agent's competence more than the positive inaccurate agent's competence. It is possible that the denial of reward is more attention-grabbing than receiving an un-earned reward. Follow-up analyses showed that age and self-esteem were not significant in predicting children's likelihood to show this effect.

It is interesting that higher age significantly predicted choosing the positive inaccurate puppet as having the easiest game. Though the modal response was that children thought all the games were easy, it appears that older children may interpret the presence or absence of rewards as signifying difficulty level in a way that younger children do not. It is possible that older children have more experience with contingent feedback and rewards, and that experience leads older children to relate rewards to good performance. Another possible explanation is that, since these questions were asked at the very end of the game, it may have been more demanding on the younger children's memory, so that they did not recall how easy or difficult they had experienced the games to be.

Differences in children's performance on the task with each of the three agents may have been difficult to capture with a brief task, showing little meaningful difference in error ratio performance across within-subject puppet conditions by the lack of significant results on these analyses, though children took a significantly higher number of card turns to complete the task with the positive inaccurate than with the accurate puppet. A more difficult task may have allowed for more variation in

performance so that the error ratio analysis could be more fruitful. Since children's performance did not vary widely among the three puppets, the modal response that the puppets' games were "all easy" may have been a reflection of children realistically assessing that they performed similarly on all the puppets' games. It should be noted again that children had a significantly higher ratio of errors while playing with the third of the three puppets, likely due to fatigue. However, there was an interaction of this effect with puppet type; specifically, children had a higher ratio of errors while playing with the positive inaccurate puppet when they played with it third. Future studies could examine if individual differences in the children's thoughts about rewards or about the positive inaccurate puppet specifically would lead to differences in the children's performances on that puppet's game. For example, children who strongly preferred receiving teddy bears may be more likely to take extra card turns with the positive inaccurate puppet in order to receive more teddy bears, and rushing through the negative puppet's game which provides no teddy bears, than children who were less interested in the incentive. Perhaps children who felt guilt or other negative emotions toward the positive puppet might also take longer on its game as the emotional experience of playing with the puppet could have caused slower performance.

In the comprehension checks, it is interesting that above 90% of the children described the behavior of the puppets as the puppets truly behaved on 5 of the 6 questions, but less than 80% of children state that the positive inaccurate puppet did not reward their non-matches, even though it did. One explanation is that the children did not accurately encode or recall the positive inaccurate puppet's behavior, so that this measure was a true comprehension check that above 20% of children failed. Receiving an un-earned reward may not have made as much of an impression on the children, so that it was not as perceptible as the inaccuracy of the negative inaccurate puppet, which withheld a reward it had initially told the children it would give, which above 90% of children stated had occurred.

It is unclear if children answered the question of the positive inaccurate puppet as an honest re-telling of what had occurred. Perhaps some children felt guilty, or feared admitting to the experimenter, that the agent rewarded them when they hadn't "earned" a reward and thus mis-stated the puppet's behavior in that instance. Children may not have deliberately lied outright, but may have stated wishful thinking rather than what actually happened, as children in experimental contexts sometimes state wishful thinking for a desired outcome (Stipek & Tannat 1984). A comprehension check asking children explicitly to state what the puppet really, truly, did, and if they thought the puppet was right or wrong (or "silly") could address any misgivings the children had about the positive inaccurate puppet, whether moral misgivings about un-earned rewards, or perception of experimenter demand that the experimenter would think it was wrong for the child to receive an un-earned reward, that might have interfered with the children's reporting of its behavior exactly as it had occurred. As the question was worded in this study, it may not be a true comprehension check, assessing what the children took away from the puppets' behavior, but instead may have given too much leeway for the child to respond in a manner of wishful thinking rather than exact reporting.

Experiment 1 assessed both individual and developmental differences in whether children endorsed the competence of an inaccurate, but favorable, interaction partner. The strength of the cross-sectional design is to compare a relatively large sample of children across age groups. One question a cross-sectional design can't answer, however, is whether the child's choice of most competent (or favorite) agent would change as a function of development over time, and whether any change in choice of agent can be predicted by the child's (change in) self-esteem or the child's age at start and end points of the study. Another weakness is that it can't be determined if the findings of Experiment 1 are truly developmental differences and aren't specific to any confounding shared characteristics of an age cohort in the cross-sectional sample, such as shared cultural characteristics or

classroom experiences that might impact the children's experiences of the agents in the study or the children's self-reported self-esteem levels.

To address the limitations of Experiment 1, Experiment 2 introduced a longitudinal element. A strength of running a longitudinal study would be that any change in the child's choice of most competent or favorite puppet could be assessed in relation to any change in the child's self-esteem, as well as change in the child's age. Another strength of running a longitudinal study would be, that if the same age difference is seen in both the cross-sectional study and the longitudinal study, it would suggest a robust developmental effect, controlled for any specific shared confounding characteristics of the cross-sectional sample.

Another strength of a longitudinal study is examining whether high self-esteem may be more predictive of endorsing the positively-valenced inaccurate puppet's competence as children become older. A child's level of self-esteem may change across the preschool years, likely becoming lower as the child moves toward elementary-school age (Robins, Trzesniewski, Tracy, Gosling, & Potter 2002). If high self-esteem correlates with a tendency to endorse positive feedback-givers, even if inaccurate, as competent, and self-esteem lowers from preschool to elementary school, high self-esteem may be more predictive due to this variation. However, it is also possible that the difference between an individual child's self-esteem in early preschool versus late preschool is a smaller difference than the mean cross-sectional age cohort difference between early preschool and late preschool. A longitudinal study can reveal whether there is significant change over time in children's choice of the most competent agent and whether that change over time depends on age at their first and second measurements.

Experiment 2

Experiment 1 examined the relative contributions of developmental and individual differences in assessing preschool-aged children's perception of the competence of agents. To extend the question of

whether the valence of an inaccurate agent influences whether children endorse that agent as competent, a longitudinal follow-up study further examined individual as well as developmental differences in the task. In Experiment 2, children who participated in the first study were followed up at an additional time point, once they had entered the next age cohort, on the agent task to track individual and developmental differences. I measured whether there was significant change over time, from time 1 to time 2, in the agent type that children chose as their favorite and most competent, and if that choice was predicted by the child's age at time 1. To date, Experiment 2 might be the first study to examine inaccurate agent judgments over the timespan of a year longitudinally.

Previous prospective studies of inaccurate agents have tested children's memory of agents by following up with children no more than one week later and found that children form enduring profiles of inaccurate agents and use those profiles when evaluating new testimony from those agents (Corriveau & Harris 2009). In contrast, Experiment 2 used a longitudinal design to examine developmental differences across a much longer time span to measure any changes in perception children have from early preschool to later preschool. Experiment 1 did not reveal significant findings for self-esteem. It is possible that self-esteem is more predictive of children's choice of most competent or favorite agent when measured at two time points, as a possible interaction with children's age, so that self-esteem may be more predictive at higher than lower ages, or as an interaction with children's choice of agent at time 1 for favorite or most competent agent.

Method

The design and procedure was identical to Experiment 1. Please see the methods section of Experiment 1. The experimenters for Experiment 2 are 1 White female and 1 Black female from experiment 1. As 24 children re-enrolled at the nursery school for the current academic year, there were a potential 24 participants to be recruited, with 16 having been age 3 at time 1 and are age 4 at

time 2, and 9 having been age 3 at time 1 and are age 5 at time 2. (These 9 were older 3-year-olds in the time 1 sample, who had aged out of the 4-year-old cohort by the time they were tested for time 2.) Of the 24 potential children, 1 refused to participate and 2 unenrolled from the school before the experimenter could test them. The final sample for Experiment 2 comprised 21 children who completed the puppets session at time 1 and time 2. Of those 21 children, 10 had completed the self-esteem session at time 1 and 18 completed the self-esteem session at time 2. Ten of the 21 children completed the self-esteem measure at both time 1 and time 2. Due to this low sample size, Experiment 2 is underpowered, particularly for the self-esteem analyses. Therefore, these underpowered results should be considered preliminary and interpreted with caution.

Design, Analysis, and Power – Experiment 2

The analyses in Experiment 2 are mostly logistic regressions on dichotomous criterion variables, whether the child chose or did not choose a particular agent to answer the given question at time 2. The predictor variables include age as a continuous predictor, total score on the Harter and Pike (1984) self-esteem scale, children's choice of agent for a given question at time 1, and interactions of these variables. These analyses will report the log odds (standardized B) as well as the odds ratios (Exp(B)). All tests are two-tailed with an alpha of 0.05. Effect size and power for these logistic regressions are low, with effect sizes below 0.3 and power below 0.2. A few analyses in Experiment 1 require t-tests, with effect sizes in the 0.2 to 0.3 range and power in the 0.50 to 0.60 range.

Results – Experiment 2

Data were analyzed in a similar manner to Experiment 1. Testing these exploratory hypotheses in the longitudinal design of Experiment 2 can provide preliminary results to guide future research designs. Attrition in longitudinal research is common, unfortunately resulting in smaller samples and therefore lower power.

Influence of favorability of feedback on whether child sees an informant as competent.

Favorite agent. The criterion variables were dichotomous, whether the child chose a given agent (accurate, positive inaccurate, or negative inaccurate) for the question of which puppet was their favorite. A paired samples t-test between time 1 and time 2 revealed a marginal but non-significant finding that, as a group, children were more likely to choose the positive inaccurate puppet as their favorite at time 2 than time 1, $t(19)=-2.042$, $p=0.055$. The paired samples t-tests for the accurate and negative inaccurate puppets were not significant. This finding mirrors the finding from experiment 1 that older children were more likely than younger children to choose the positive inaccurate puppet as their favorite.

Three logistic regressions examined whether the dichotomous criterion variable (accurate, positive, or negative inaccurate agent as most competent at time 2)was predicted by choice of favorite agent at time 1, child's mean-centered age at time 1, and the interaction between the child's mean-centered age at time 1 and choice of favorite agent at time 1. Logistic regression analyses did not reveal significant findings for any of these predictors for any of the three logistic regressions: accurate, positive inaccurate, or negative inaccurate.

Most competent agent. The criterion variables were dichotomous, whether the child chose a given agent (accurate, positive inaccurate, or negative inaccurate) for the question of which puppet was most competent. A paired samples t-test between time 1 and time 2 revealed that, as a group, children were more likely to choose the accurate puppet as most competent at time 2 than time 1, $t(19)=-2.333$, $p<0.05$. The paired samples t-tests for the positive and negative inaccurate puppets were not significant. Three logistic regressions examined whether the dichotomous criterion variable (accurate, positive, or negative inaccurate agent as most competent at time 2) was predicted by choice of most competent agent at time 1, child's mean-centered age at time 1, and the interaction between the child's mean-

centered age at time 1 and choice of most competent agent at time 1.. Logistic regression analyses did not reveal significant findings for any of these predictors for any of the three logistic regressions: accurate, positive inaccurate, or negative inaccurate.

Self-esteem as a moderator of whether favorability of feedback influences whether children see the informant as competent.

Self-esteem. Children's mean self-esteem scores and standard deviations in Experiment 2 were comparable to the values found in Harter and Pike (1984) when the scale was originally normed, suggesting that there was sufficient variation in this sample to detect effects. Please see Table 1 with descriptive statistics for the self-esteem measure in Experiment 2. The hypothesis was that self-esteem might moderate the effect of valence manipulation on children's choice of competent puppet and favorite puppet; specifically, the hypothesis was that children with high self-esteem would choose the positive inaccurate puppet as most competent and most preferred agent. Longitudinally, I predicted that self-esteem may stay relatively stable from time 1 to time 2, though could potentially go down, perhaps reflecting that children become more calibrated in assessing their abilities as they grow older. The results showed that self-esteem was not significantly different between time 1 and time 2, paired samples t-test, $t(9)=-1.191$, $p=0.264$. Because self-esteem did not differ significantly from time 1 to time 2, I created an average self-esteem score across time points for each participant, and then mean-centered that average for each participant.

It should be noted that only 10 children completed the self-esteem measure at both time 1 and time 2, so these analyses are underpowered and likely insensitive to detect real effects. I ran 6 logistic regressions to examine whether informant type (accurate, positive inaccurate, negative inaccurate) chosen for favorite and most competent at time 2 was predicted by mean-centered averaged self-esteem (abbreviated "SE"), the two-way interaction of mean-centered age at time 1 by SE, the two-way

interaction of choice at time 1 by SE, and the three-way interaction of mean-centered age at time 1 by puppet choice at time 1 by SE. None of these regression analyses were significant. I was able to run 5 of the 6 analyses successfully because too few children chose the negative inaccurate puppet as most competent at time 2 to be able to run that analysis. For the accurate puppet as favorite and accurate puppet as most competent at time 2, the omnibus test was significant, but none of the predictors were significant, and the intercept was not significant.

General Discussion

Summary and Explanation of Expected Results

Experiments 1 and 2 presented preschool children with interaction partners who provided or denied rewards accurately or inaccurately to children while they played a game. Children's verbal responses rating their preference for and competence of the agents, verbal ratings of how well they themselves performed on the game, and verbal responses of their memory of the agents' actions were assessed. Additionally, behavioral responses of whether children were more likely to follow the lead of an inaccurate informant by playing the game according to how the inaccurate agent acts, or how the children have been previously instructed is the correct way to play the game, were assessed.

The conceptual motivation for Experiments 1 and 2 was to assess whether the favorability of feedback influences which agent children choose as a favorite, and whether young children endorse the feedback-giver as competent, and whether self-esteem moderates these effects such that young children with high self-esteem more likely to choose as a favorite and endorse as competent an inaccurate agent who provides favorable feedback. The experiments aimed to examine developmental differences in these findings cross-sectionally as well as examining whether children's choice of agent as favorite and most competent over time, and whether it is predicted by children's age and choice of puppet at time 1 in a longitudinal study. The expected results were that young children with high self-

esteem would be more likely to choose as a favorite and endorse as more competent a positive, but inaccurate, feedback-giver, similarly to adults (Kernis & Sun 1994, Taylor & Brown 1988). Additionally, younger children were expected to endorse a positive, but inaccurate, feedback-giver as more competent, as compared to older children, as younger children's person perception may be more affected by social-emotional information than competence information (Brosseau-Liard & Birch 2010, Lane, Wellman, & Gelman 2013, Cain, Heyman, & Walker 1997, Stipek & Daniels 1990), and younger children take positive information about their own performance into account more than negative information (Stipek, Roberts, & Sanborn 1984, Bjorklund & Green 1992).

There was a lack of support for the hypotheses, and, in fact, evidence for the opposite developmental conclusion, that higher age predicts choosing the positive inaccurate puppet as favorite. Self-esteem did not appear to influence children's judgments of most competent and favorite agent. This may be due to low statistical power. Alternatively, it's possible that null self-esteem effects are due to the measure itself and the relevance of the questions to the participants. Self-esteem may not relate to children's judgments of the most competent and favorite agent, because these judgments might be influenced more by other individual differences such as in executive function, theory of mind, or personal experiences with rewards and feedback. Self-esteem may not be a relevant construct in children's assessment of the favorability of inaccurate feedback or the competence of those who provide inaccurate, even if favorable, feedback. It is interesting that higher age, rather than lower age, predicted endorsing the positive inaccurate puppet as the favorite in experiment 1. This finding was replicated in experiment 2 in that children were more likely to choose the positive inaccurate puppet as the favorite at time 2 than at time 1. As stated in the discussion section of experiment 1, it is possible that children have more experience with feedback as they grow older, and so have learned from experience to value positive feedback, perhaps even to the point of valuing inaccurate positive

feedback. The longitudinal finding suggests that this is a developmental phenomenon and not merely an individual differences phenomenon.

One interesting feature of this study is that children played with three puppets sequentially and made judgments about all three puppets. All children played with the accurate puppet first, who established a baseline rule of awarding teddy bears for matches and not awarding teddy bears for non-matches. Presumably, children then made decisions about the negative inaccurate and positive inaccurate agents relative to the accurate puppet. The order effect for error ratio of card-pair turns the child made when playing with the positive inaccurate puppet second (fewer errors) versus third (more errors) may suggest that children are comparing the positive and negative inaccurate puppets against each other differently when the child sees the positive inaccurate puppet second versus third. Children may be especially sensitive to contingent interactions with the positive versus negative inaccurate puppets, but fatigue may heighten these effects of seeing the positive inaccurate puppet last. Or, the order effect might matter that when children see the negative inaccurate puppet second rather than third, this may lead children to be, in a way, more desiring, for teddy bears from the third puppet, and they may be "playing the system" after the deprivation of rewards to gain more teddy bears from the puppet that always rewards, and therefore make more errors. This experiment may have ecological validity in that children may be treated one way by one adult, but then held to a different set of rules by another adult, which may be a confusing experience at the preschool age.

There is a question of whether the children are interpreting the inaccurate teddy bears as incompetent or instead as socially motivated, such as the positive inaccurate puppet acting "nice" or "silly" and the negative inaccurate puppet acting "mean" or "cheating." If children are interpreting the puppets this way, it is a confound of the study. Children may then generalize the "niceness" or "meanness" of the puppet to the puppet's competence, as in previous studies. Future studies could

specifically ask the children if they are viewing the puppets as nice or mean, competent or incompetent. Though this study asks the child which puppet is their favorite, and presumably the child would take any social emotional evaluation of the puppet into a holistic evaluation when deciding on their favorite puppet, it is a rather indirect way of addressing this question. Asking the child to rate the puppet's social-emotional qualities directly, their "warmth" and not only their "competence", would better address this question.

Broader Implications

If young children with particularly high or low self-esteem are more affected by the emotional content of an interpersonal evaluation to the point that they interpret feedback less objectively, this may have negative implications for their learning. Studying the relationship between self-esteem and interpretation of feedback at different ages can tell us about whether these effects may compound as children grow older and experience more varied feedback, or if these effects might be heightened for some developmental stages but not others. As this study did not provide evidence for a relationship between self-esteem and responding differently to the valence of inaccurate information, this study can't conclude that children with high self-esteem are any less objective in interpreting inaccurate feedback than are children with low self-esteem, or vice versa.

Children begin to build their self-schemas at a very young age. A high, even unrealistically high, level of self-confidence in early childhood may be adaptive (Bjorklund & Green 1992), as it might motivate children's exploration and practice of tasks that they might not otherwise seek out if they had a more realistic grasp of their abilities. Therefore, it may be adaptive at this age for children to pay more attention to the emotional content of feedback than its factual content. This study, while not answering the question of adaptiveness, can provide information as to whether children are especially affected by the valence of feedback when the feedback does not accurately describe their performance.

It appears that older children are more sensitive to positive inaccurate feedback than are younger children. It is possible that by the end of the preschool years, children are becoming more calibrated in perceiving differences in valence among those who give them feedback. Perhaps younger children are more oblivious to feedback than older children, or do not derive the same meaning from feedback as do older children.

By the time a child reaches adulthood, he or she will have had experience with thousands of examples of interpersonal feedback in the context of learning. While high self-esteem, and its corresponding tendency to endorse positive feedback about oneself as correct, may or may not be adaptive to an adult, particularly inflated self-esteem, as in narcissism (Kernis & Sun 1994), or particularly low self-esteem, as in depression (Taylor & Brown 1988), may be problematic to personal growth and career development in adults. Identifying the early emerging profiles of children with high or low self-esteem and how they interpret feedback could allow for early childhood interventions to stop any maladaptive self-schemas before individuals with maladaptive self-schemas practice these associations over the course of many years before adulthood. This study does not provide evidence for the relationship between self-esteem and reaction to feedback in this age range, but it does provide developmental evidence that older children are more sensitive to positive inaccurate feedback than are younger children.

Additionally, this research has implications for early childhood education. While this work cannot directly address current debates about the academic content of preschool curriculums, it could provide information about how children of these ages interpret the relational and emotional content of feedback at this age. Older children choose an always-rewarding agent as their favorite, even if the agent occasionally rewarded them incorrectly, while younger children did not show this pattern. This study does provide some evidence that older children think activities for which they received constant

positive feedback, even if incorrect, are "easier" while younger children do not show this pattern. It is possible that, in evaluating the difficulty level of the task, older children are focusing on the rewards they received in addition to, or perhaps sometimes more than, recalling the effort they expended to achieve the rewards. A child's sense of which task is "easier" may not translate into their actual performance on the task, and so early childhood educators may want to encourage children to focus on the effort involved in bringing about an outcome and consider whether using rewards may undermine this focus. Further studies of whether children at this age believing a task is "easier" lead to better performance, or at least more exploration, of a novel task could provide more evidence to make recommendations to early childhood educators.

Children's tendency to follow the lead of a positive inaccurate agent and question a negative inaccurate agent could have implications for how persuadable children are in responding to media and advertising. Though children at this age can identify an advertisement (McAlister & Cornwell 2009), children could be taught to be more skeptical of friendly and congenial characters who try to sell them products in advertisements. This study found evidence that children were more likely to question incorrect feedback when the feedback was not in their favor, in terms of placing their cards in the game where they thought they belonged, rather than where the agent incorrectly said they did, and more likely to go along with incorrect feedback that was in their favor, by placing the cards where the agent incorrectly said they should, rather than where they truly thought. Though media may not tell the child anything directly factually incorrect, perhaps teaching children to proceed more critically could be one way to counteract this susceptibility to questioning a congenial agent's actions less.

Strengths and Limitations of the Current Research

The current study's strengths include manipulating both the valence and accuracy of feedback provided by an agent, so that children perceive multiple traits of the agent at once in deciding on his or

her favorability and competence. The current study also has the strength of collecting longitudinal data to determine if any moderating effect of self-esteem on whether children are affected by the valence of an agent's feedback information in deciding on a preference for him or her, and his or her competence is a robust developmental effect. The combination of cross-sectional and longitudinal studies provides aggregated evidence that self-esteem may not relate to interpreting differently-valenced incorrect feedback, albeit with the strong caveat that low power may have prevented detection of such effects.

The current study has a number of limitations. First, both Experiment 1 and Experiment 2 are underpowered. A larger sample size could determine if any lack of significant findings may be due to inadequate sample size. Second, the study used a within-subjects design that did not fully cross valence and accuracy in a 2x2 design. While it might be less ecologically realistic to produce an informant that is negatively valenced and accurate for instances where the child is correct, and care would need to be taken to make incompetence clear in all conditions so that children doesn't assume that they are simply playing with two puppets who play by different rules, a fully crossed 2x2 design would allow for direct comparisons between conditions without the possible issue of taxing the child's memory for agents or experimental fatigue. Third, the study is limited by procedural constraints requiring the self-esteem measure and the feedback agents' task to be run in separate sessions, while running both sessions in the same day in a future study would be preferable to gain a more fine-grained assessment of the child's development at that point in time and to prevent attrition. Fourth, it is unclear if the children are interpreting the puppets' instructions in terms of the puppets' competence or social-emotional qualities. Future studies could explicitly ask the children for their ratings of both the puppets' competence and social-emotional qualities.

Additional Directions for Future Research

Given cultural variation in how children are encouraged to construct and express self-views (Wang & Ollendick 2001, Zayas & Solari 1994), this review which examines available literature, and the current research project which examines self-views of and reactions to feedback by participants attending nursery school in an affluent community, should be considered as describing these processes in a Western, middle-to-high income sample, of varying races and ethnicities though predominantly White-American and its conclusions should be applied cautiously beyond this population. Further research is needed to characterize these processes more broadly. Several studies have documented cultural variation in encouraging independent versus interdependent construals of self in adults and children (Markus & Kitayama 1991). Independent cultures encourage self-advocacy and viewing the self as unique. Interdependent cultures encourage viewing the self in relation to others and accommodating others' needs. North American and Western European cultures are primarily independent and East Asian and South Asian cultures are primarily interdependent, so future studies could examine Asian versus European or North American populations. It is possible that children with an interdependent view of self may be more negatively affected by receiving or being denied a reward inaccurately due to the disharmonious interpersonal context it creates, or children with an interdependent view may rate their own performance more in terms of the behavior of the agent to maintain a harmonious social relationship. SES also relates to independence versus interdependence; high-SES culture encourages independence while low-SES culture encourages interdependence (Snibbe & Markus 2005).

This study was exploratory in examining the relationship between an individual-difference measure, self-esteem, and children's perceptions of competence of agents who provide feedback. Other individual differences constructs aside from self-esteem should be explored and may be more relevant. First, some differences in children's responses to the positive-valenced inaccurate puppet may be due to a lack of inhibitory control, and executive function measures examining inhibitory control

could be useful in ruling out this possible explanation for children's responses. Theory of mind may be another useful construct to examine. Young children may not fully understand what it means for an agent in this task to misunderstand what matching or non-matching shapes are; this knowledge may seem so obvious to children that they may attribute the agent's actions not to competence but to a social-emotional motivation, such as wanting to be nice or mean. Directly asking children why they think the agent responded the way it did, to ascertain whether children are thinking of competence or "nice" or "mean" motivations, would be useful. In combination with a theory of mind measure, it could provide interesting information about children's attributions of motivations to others in ambiguous contexts.

Another way to study children's ability to view feedback objectively would be for children to rate the agents' responses to another child. It is likely that children may be more objective in assessing another child's performance and objective in assessing the accuracy or inaccuracy of an agent's feedback for another child more than for themselves. Children could play the task similarly to Experiment 1 and 2, but also view another child playing the task and rate the child's performance and the competence of the puppets' rewarding patterns. Such a task could determine children's level of objectivity when the valence of information is crossed with inaccuracy.

Concluding Comment

These studies aimed to provide information of children's relative weighting of competence and favorability of information in a dynamic interaction. While previous studies have examined children's relative weighting of valence of social-emotional and competence information from static descriptions of agents, these studies provide valence and accuracy manipulated separately in the same live interaction. This study provided evidence that at older ages, children are more likely to favor positive feedback, even if that feedback is inaccurate. Additionally, these studies add to the inaccurate

informant literature by contributing information about children's sensitivity to accuracy based on the valence of information. These studies aimed to describe children's behavior toward rejecting or accepting feedback that is inaccurate, even if in their favor, during the preschool years, and do not provide evidence that self-esteem moderates children's judgments of the competence or favorability of inaccurate informants who provide positively-valenced or negatively-valenced information. Finally, these studies examine children's sensitivity at different ages to the favorability of feedback, even if it's incorrect.

References

Bjorklund, D. F., & Green, B. L. (1992). The adaptive nature of cognitive immaturity. *American Psychologist, 47(1),* 46.

Brosseau-Liard, P., & Birch, S. (2010). "I bet you know more and are nicer too!" What children infer from others' accuracy. *Developmental Science, 13(5),* 772-778.

Brown, J.D. (1986). Evaluations of self and others: Self-enhancement biases in social judgments. Social Cognition, 4(4), 353-376.

Burhans, K. K., & Dweck, C. S. (1995). Helplessness in early childhood: The role of contingent worth. *Child development, 66(6),* 1719-1738.

Cain, K.M., Heyman, G.D., & Walker, M.E. (1997). Preschoolers' ability to make dispositional predictions within and across domains. *Social Development, 6,* 53-75.

Chao, C., Harter, S., Adams, P., & Strop, J. (1983). Dimensions underlying children's perceptions of competence. Unpublished data, University of Denver.

Cohen, J. (1992). A power primer. *Psychological bulletin, 112(1),* 155.

Corriveau, K.H., Meints, K., & Harris, P.L. (2009). Early tracking of informant accuracy and inaccuracy. *British Journal of Developmental Psychology, 27(2),* 331-342.

Dawson, E., Gilovich, T., & Regan, D. (2007). Motivated reasoning and susceptibility to the ''Cell A'' bias. Retrieved from

http://www.researchgate.net/publication/237429074_Motivated_Reasoning_and_Susceptibility_to_the_Cell_A_Bias.

Davis-Kean, P. E., & Sandler, H. M. (2001). A meta-analysis of measures of self-esteem for young children: A framework for future measures. *Child development, 72(3),* 887-906.

Dodge, K. A., & Tomlin, A. M. (1987). Utilization of self-schemas as a mechanism of interpretational bias in aggressive children. *Social cognition, 5(3),* 280-300.

Fantuzzo, J. W., McDermott, P. A., Manz, P. H., Hampton, V. R., & Burdick, N. A. (1996). The Pictorial Scale of Perceived Competence and Social Acceptance: Does It Work with Low-Income Urban Children?. *Child Development, 67(3),* 1071-1084.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods, 39(2),* 175-191.

Johnson, S. C., Dweck, C. S., & Chen, F. S. (2007). Evidence for infants' internal working models of attachment. *Psychological Science, 18(6),* 501-502.

Johnson, S. C., Dweck, C. S., Chen, F. S., Stern, H. L., Ok, S. J., & Barth, M. (2010). At the intersection of social and cognitive development: Internal working models of attachment in infancy. *Cognitive Science, 34(5),* 807-825.

Harter, S., & Pike, R. (1984). The pictorial scale of perceived competence and social acceptance for

young children. *Child development,* 1969-1982.

Harter, S. & Wright, K. (1984). The relationship between social support, depression, and young children's perceptions of competence and social acceptance. Unpublished manuscript, University of Denver.

Kernis, M. H., & Sun, C. R. (1994). Narcissism and reactions to interpersonal feedback. *Journal of Research in Personality, 28(1),* 4-13.

Kim, S., Kalish, C. W., Weisman, K., Johnson, M. V., & Shutts, K. (2015) Young children choose to inform previously knowledgeable others. Manuscript in press at the *Journal of Cognition and Development*.

Koenig, M.A., & Doebel, S. (2013) Children's understanding of unreliability: Evidence for a negativity bias. In Banaji, M. R., & Gelman, S. A. (Eds.). *Navigating the social world: What infants, children, and other species can teach us*. (pp 235-240). Oxford University Press.

Koenig, M.A., & Jaswal, V.K. (2011). Characterizing children's expectations about expertise and incompetence: Halo or pitchfork effects? *Child Development, 82(5),* 1634-1647.

Lane, J. D., Wellman, H. M., & Gelman, S. A. (2013). Informants' traits weigh heavily in young children's trust in testimony and in their epistemic inferences. *Child development, 84(4),* 1253-1268.

Langer, E.J. (1975). The illusion of control. *Journal of Personality and Social Psychology, 32(2),* 311.

MacDonald, K., Schug, M., Chase, E., & Barth, H. (2013). My people, right or wrong? Minimal group

membership disrupts preschoolers' selective trust. *Cognitive Development, 28(3),* 247-259.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. Psychological review, 98(2), 224.

McAlister, A. R., & Cornwell, T. B. (2009). Preschool children's persuasion knowledge: The contribution of theory of mind. *Journal of Public Policy & Marketing, 28(2),* 175-185.

Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological Bulletin, 130(5),* 711.

Nurmsoo, E., & Robinson, E.J. (2009). Children's trust in previously inaccurate informants who were well or poorly informed: When past errors can be excused. *Child Development, 80(1),* 23-27.

Pasquini, E., Corriveau, K., Koenig, M., & Harris, P. (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology, 43(5),* 1216-1226.

Robins, R. W., Trzesniewski, K. H., Tracy, J. L., Gosling, S. D., & Potter, J. (2002). Global self-esteem across the life span. *Psychology and Aging, 17(3),* 423.

Robins, R. W., & Trzesniewski, K. H. (2005). Self-esteem development across the lifespan. *Current Directions in Psychological Science, 14(3),* 158-162

Sabbagh, M., & Shafman, D. (2009). How children block learning from ignorant speakers. *Cognition, 112(3),* 415-422.

Sedikides, C., & Gregg, A. P. (2008). Self-enhancement: Food for thought. *Perspectives on Psychological Science, 3(2),* 102-116.

Snibbe, A. C., & Markus, H. R. (2005). You can't always get what you want: educational attainment, agency, and choice. *Journal of personality and social psychology*, *88*(4), 703.

Stipek, D. J., & Daniels, D. H. (1990). Children's use of dispositional attributions in predicting the performance and behavior of classmates. *Journal of Applied Developmental Psychology, 11(1),* 13-28.

Stipek, D.J, & MacIver, D. (1989). Developmental change in children's assessment of intellectual competence. *Child Development, 60,* 521–538

Stipek, D., Recchia, S., McClintic, S., & Lewis, M. (1992). Self-evaluation in young children. *Monographs of the society for research in child development, 57(1),* 100.

Stipek, D. J., Roberts, T. A., & Sanborn, M. E. (1984). Preschool-age children's performance expectations for themselves and another child as a function of the incentive value of success and the salience of past performance. *Child Development*, 1983-1989.

Stipek, D. J., & Tannatt, L. M. (1984). Children's judgments of their own and their peers' academic competence. *Journal of Educational Psychology, 76(1),* 75.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica, 47(2),* 143-148.

Taylor, S.E., & Brown, J.D. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychological Bulletin, 103(2),* 193.

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: the negativity bias in social-emotional development. *Psychological Bulletin, 134(3),* 383.

Wang, Y., & Ollendick, T. H. (2001). A cross-cultural and developmental analysis of self-esteem in Chinese and Western children. *Clinical Child and Family Psychology Review, 4(3),* 253-271.

Zayas, L. H., & Solari, F. (1994). Early childhood socialization in Hispanic families: context, culture, and practice implications. *Professional Psychology: Research and Practice, 25(3),* 200.

Table 1

Descriptive statistics for self-esteem scale, Experiment 1 ("T1"), n=60 and Experiment 2 ("T2"), n=10

	Mean		Standard Deviation		Variance	
	T1	T2	T1	T2	T1	T2
Cognitive	3.5	3.8	0.5	0.32	0.248	0.105
Physical	3.1	3.2	0.66	0.65	0.431	0.426
Peer	3.4	3.5	0.45	0.41	0.204	0.172
Maternal	3.3	3.2	0.57	0.53	0.322	0.281
Total score	76	73.7	9.16	10.924	83.96	119.34

Note. Subscale averages out of a possible score of 1 through 4; total averages out of a possible 23 through 92.

Table 2

Percentage of children who described each agent's behavior reflecting how the agent actually performed during the game, immediately after each agent's game and at very end of session

Question	Immediate	End of session
Accurate agent – matched pair	100	94
Accurate agent – non-match	93	88
Positive inaccurate agent – matched pair	99	86
Positive inaccurate agent – non-match	75**	59**
Negative inaccurate agent – matched pair	94	72
Negative inaccurate agent – non-match	99	75

*** p<.01*

There was also a significant overall effect of time point, immediate vs. end of session, p<.01.

Figure 1.

Experimental setup.

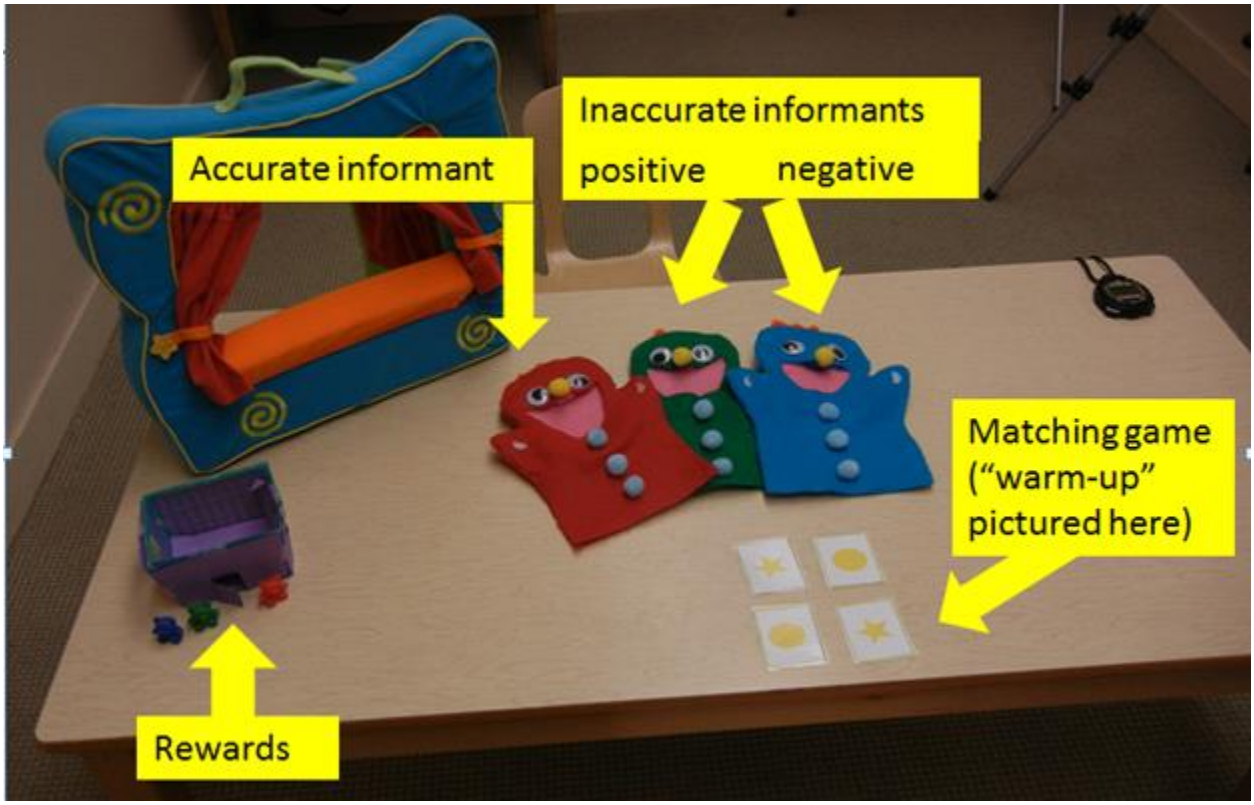

Figure 2.

Agent feedback task design.

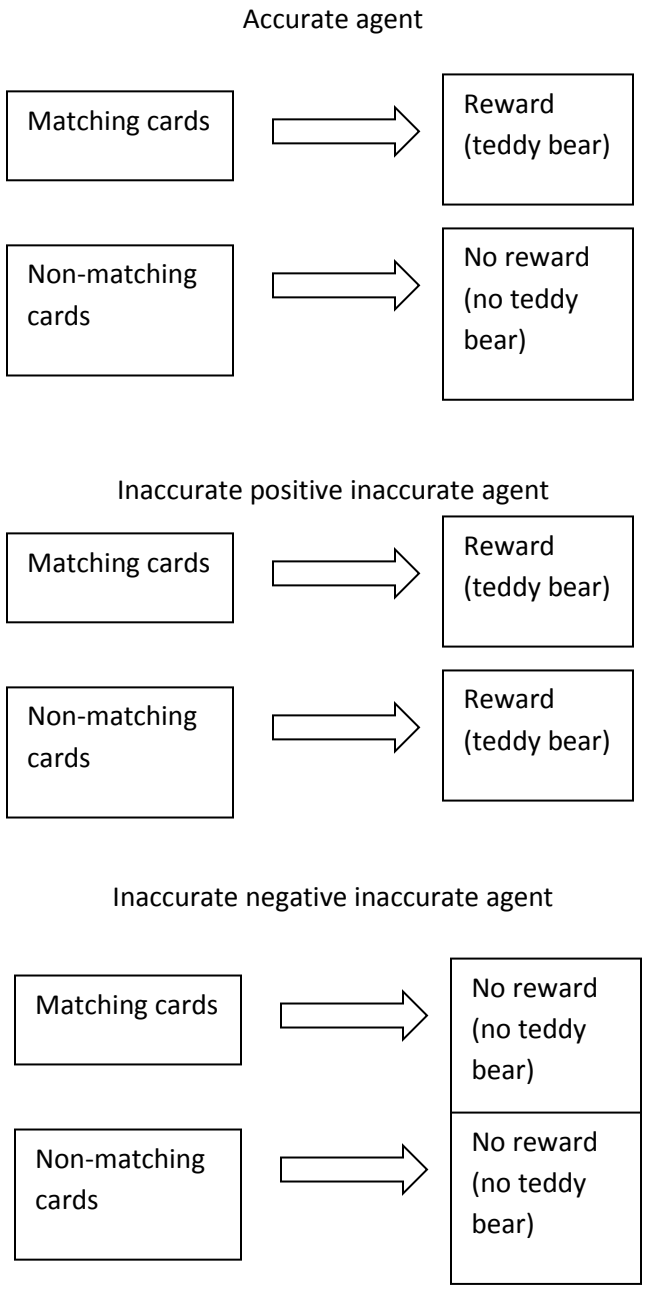

Figure 3.

Full experimental procedure.

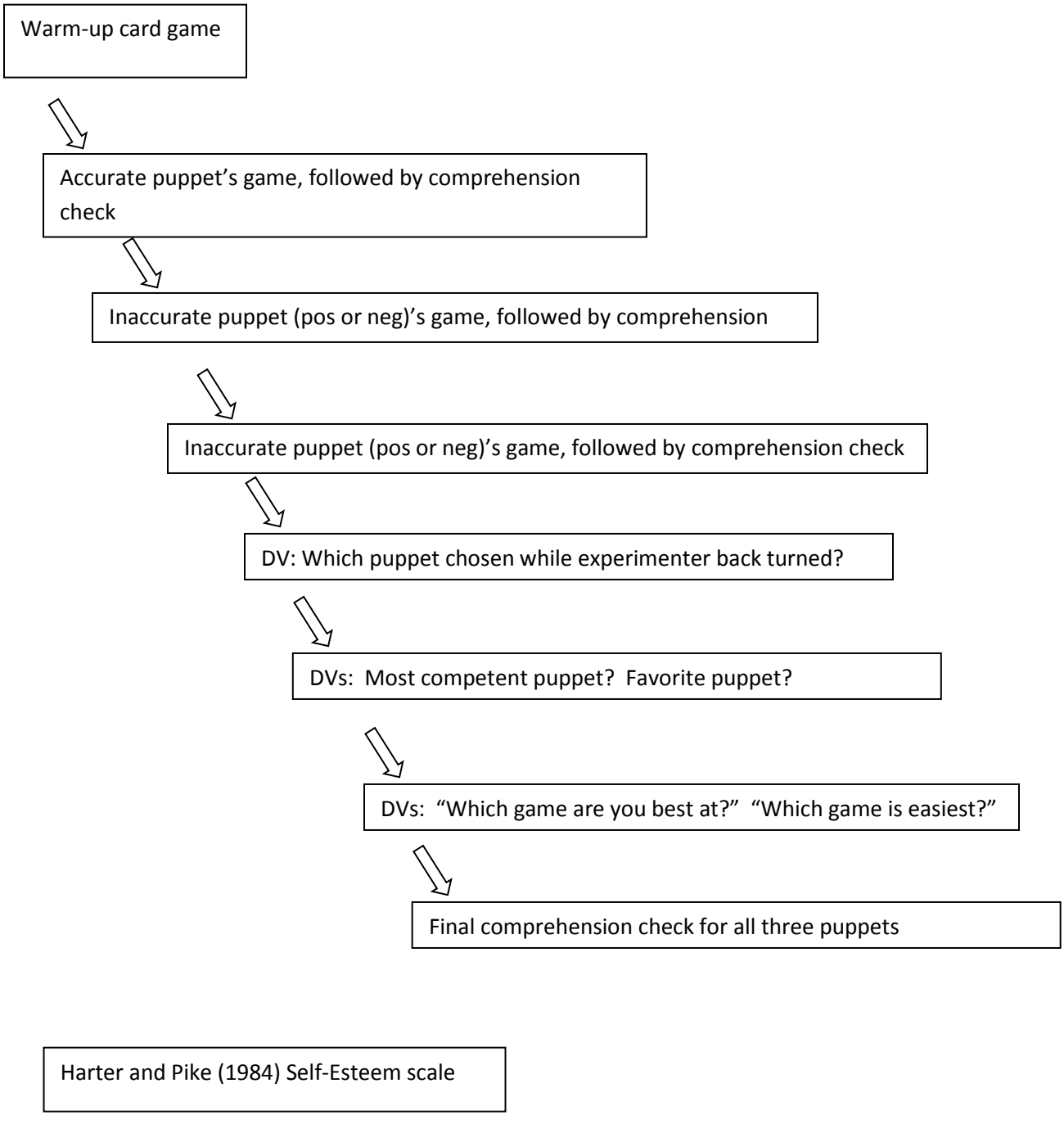

Figure 4.

Experiment 1. Results for overall sample for question, “Which puppet was your favorite to play the cards and bears game with?” Children were allowed to choose more than one agent.

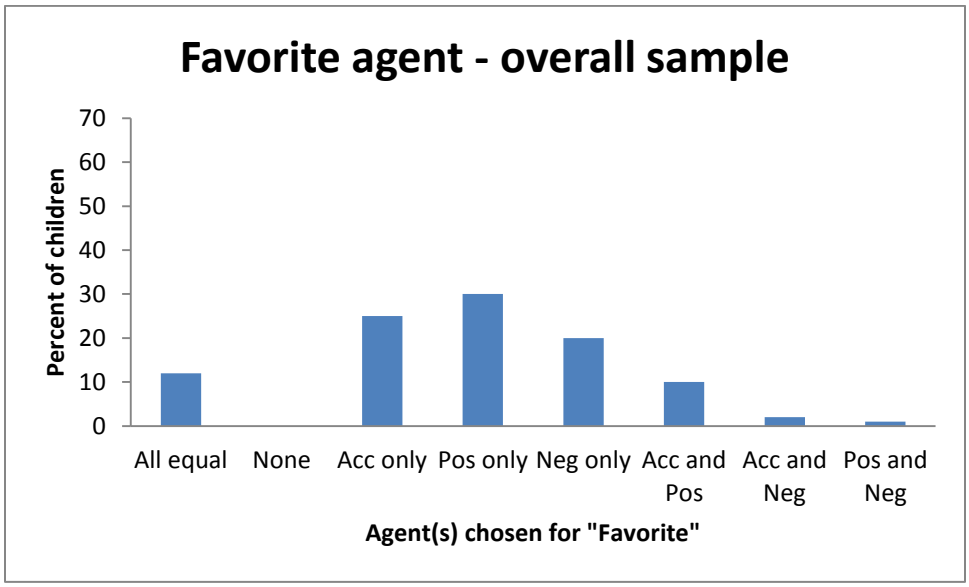

Figure 5.

Results by age for question, “Which puppet is your favorite to play the card and bears game with?” Children were allowed to choose more than one agent.

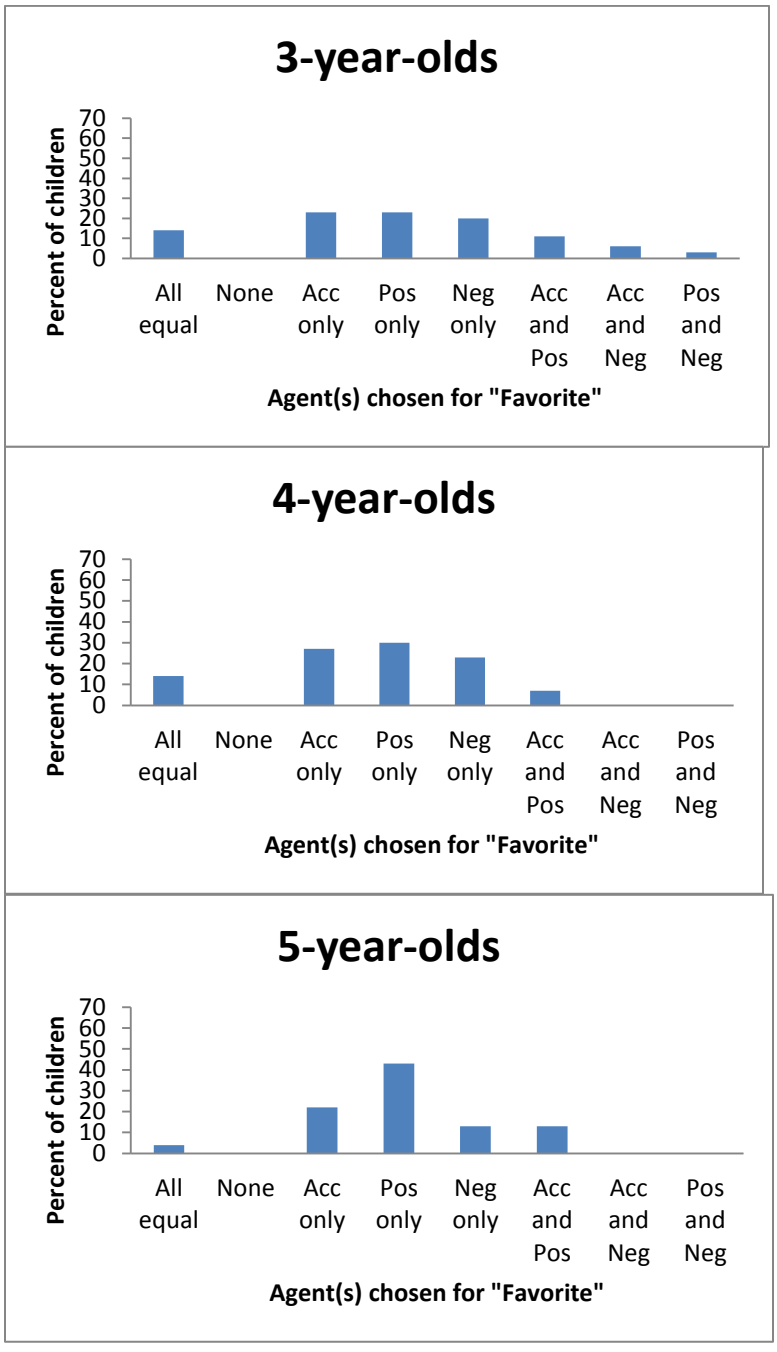

Figure 6.

Results, for subset of children who only chose 1 puppet, to the question, “Which puppet is your favorite to play cards and bears with?”

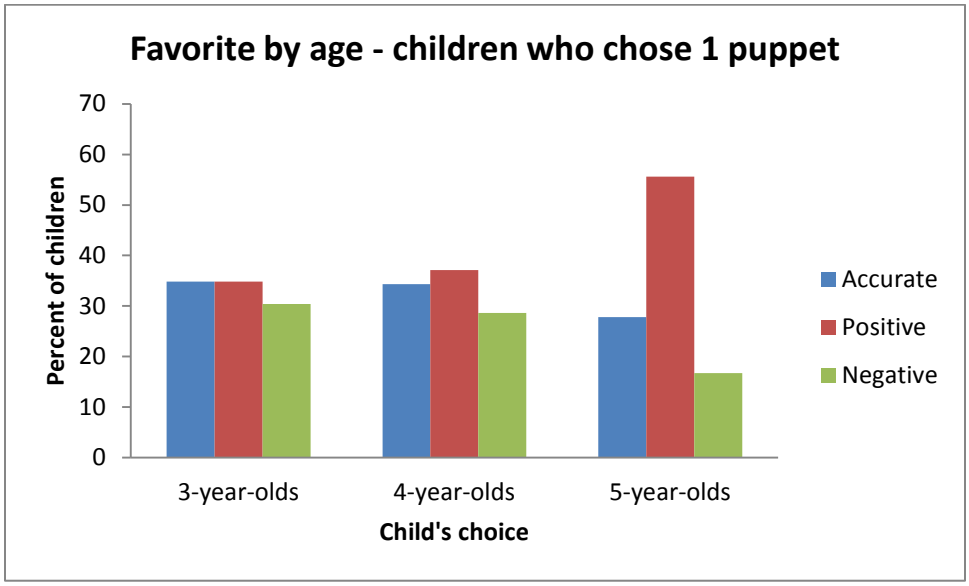

Figure 7.

Experiment 1. Results for overall sample for question, “Which puppet knows the most about how to play the cards and bears game?” Children were allowed to choose more than one agent.

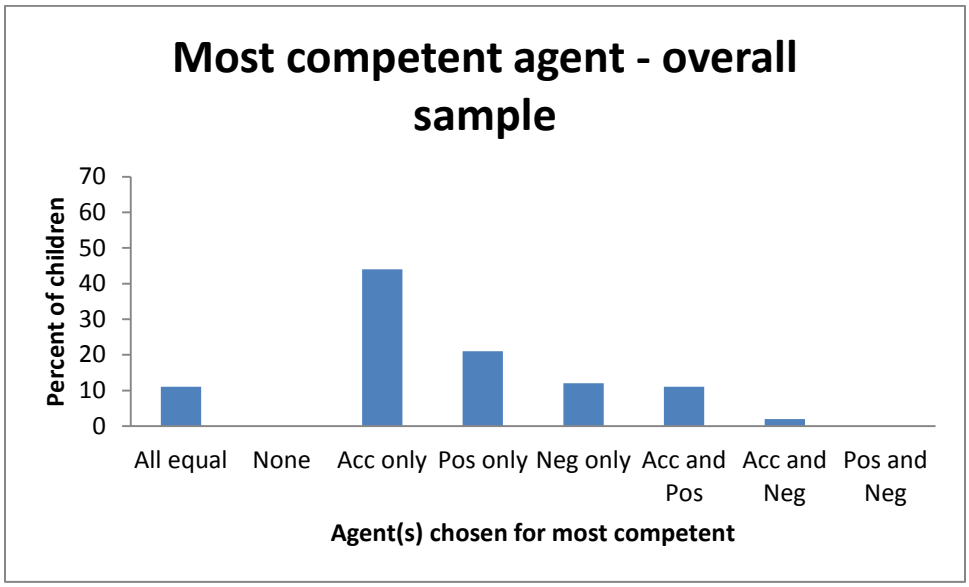

Figure 8.

Results by age for question, “Which puppet knows the most about how to play the cards and bears game?” Children were allowed to choose more than one agent.

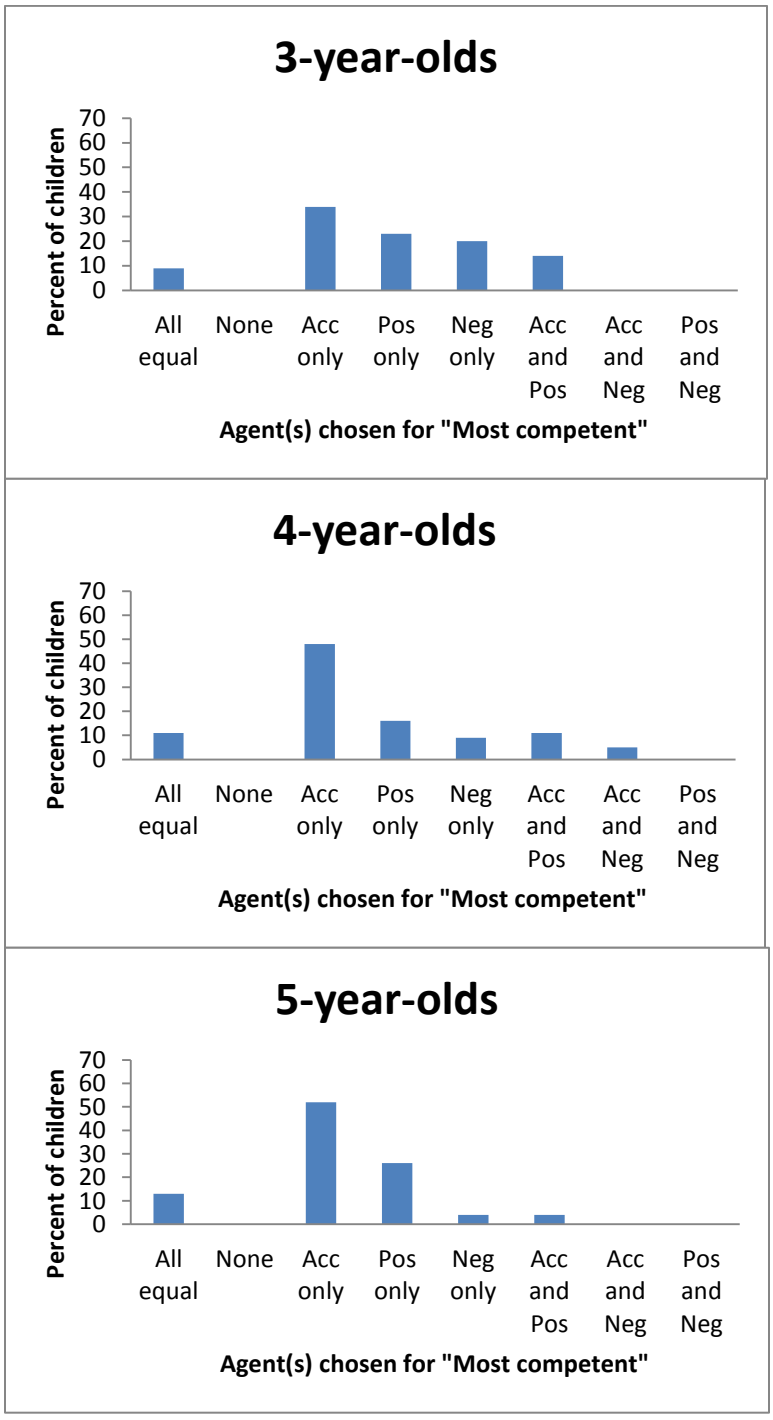

Figure 9.

Results, for subset of children who only chose 1 puppet, to the question, “Which puppet knows the most about how to play cards and bears?”

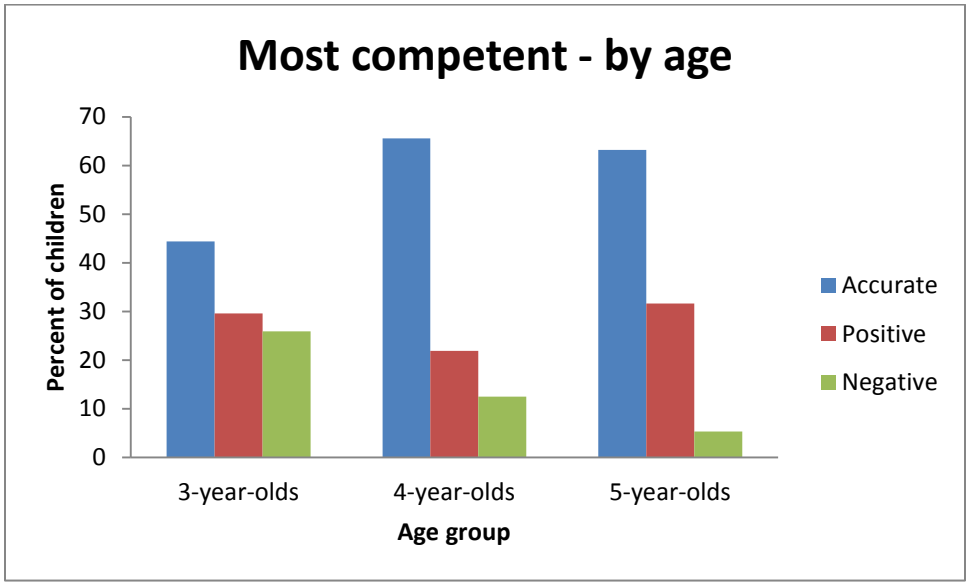

Figure 10.

Experiment 1. Results for overall sample, percent placing cards where cards actually belonged vs. where puppet incorrectly stated cards belonged.