Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Signature:

Sara Valencia Botto 6/12/2016

Infant's Emerging sensitivity to Others' Evaluation

By

Sara Valencia Botto
B.A., University of Georgia, 2014

Psychology

Philippe Rochat, Ph.D.
Advisor

Patricia Brennan, Ph.D.
Committee Member

Robyn Fivush, Ph.D.
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D. Dean of the James T. Laney School of Graduate Studies

_____
Date

Infant's Emerging Sensitivity to Others' Evaluation

By

Sara Valencia Botto
B.A., University of Georgia, 2014

Advisor: Philippe Rochat, Ph.D.

An abstract of
A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts
in Psychology
2016

Abstract

Infant's Emerging Sensitivity to Others' Evaluation
By: Sara Valencia Botto

By four years of age, children, like adults, strategically modify their behavior in the presence of an observer as a means for self-presentation. While this is evidence of an evaluative audience perception (i.e., sensitivity to other's potential social evaluation), the ontogeny of this phenomenon remains underspecified. Two studies capture the emergence of an evaluative audience perception in late infancy. In a first study 14-24 month old infants (N=49) are shown to display differential engagement towards a novel toy, as well as enhanced expression of embarrassment when the experimenter was attentive toward them. Passing as well as the way the child passed the classic Mirror Mark Test predicted such audience effect. In a second study 20 month-old infants (N=51), were tested in a situation where the Experimenter previously modelled both a positive and a negative outcome on a mechanical toy via a remote control device. Results show that infants become strategic in choosing to reproduce more positive outcomes when the Experimenter is attentive as opposed to inattentive toward them. Controlling for age, passing the Mirror Mark Test did not predict such self-evaluative audience perception. We interpret these data as confirming that an evaluative audience perception emerges by 24 months. Results are discussed in relation to an emergent self-concept that is heavily shaped by how others perceive us.

Infant's Emerging Sensitivity to Others' Evaluation

By

Sara Valencia Botto
B.A., University of Georgia, 2014

Advisor: Philippe Rochat, Ph.D.

A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts
In Psychology
2016

TABLE OF CONTENTS

INTRODUCTION 3

Self-Evaluation in Infancy. 5

The Current Study. 9

STUDY 1. 11

METHOD. 11

Participants. 11

Materials. 12

Procedure. 13

Coding 14

RESULTS. 15

Robot Task. 15

Emotion. 16

MirrorMark Test. 16

DISCUSSION 17

STUDY 2. 20

METHOD. 20

Participants. 20

Materials. 20

Procedure. 20

Coding 22

RESULTS. 22

Robot Task. 22

Temperament. 23

Emotion. 23

DISCUSSION 23

GENERAL DISCUSSION 25

CONCLUSION 28

REFERENCES 30
FIGURE 1a 34
FIGURE 1b 35
FIGURE 1c 36
FIGURE 2a 37
FIGURE 3a 38
FIGURE 3b 38

Infants' Emerging Sensitivity to Other's Evaluation

Humans are a self-conscious species; our concern with reputation management and the desire to be accepted makes us uniquely sensitive to how others perceive and evaluate us (Engelman, Herman, & Tomasello, 2012; Rochat, 2009; Rochat 2013). This sensitivity to others' potential evaluation is often manifested in our propensity to tailor our behavior as a way to maximize self-presentation in the presence of others, a phenomenon that is manifested from early childhood and throughout adulthood (Walting & Banerjee, 2007; Schlenker and Leary, 1982). For example, five-year olds as well as adults tend to be significantly more generous when others are watching (Dana, Weber, Han, 2007; Engelmann, Hamrann, and Tomasello, 2012), as well as demonstrate enhanced performance in the presence of others (Triplett, 1898; Cottrell, Wack, Sekerak, Rittle, 1968; Chevalier et al., 2012). Classic studies on strong conformity in both adult (Asch, 1956; Charrand & Bargh, 1999) and child participants (Haun and Tomasello, 2011), show the human propensity to conform to a majority opinion, even if this majority opinion or judgment is blatantly wrong or untrue.

These robust behavior modifications in the presence of others recorded in the literature paint a unique human characteristic that is prominent throughout the lifespan; we come to perceive others as potential evaluators of the self, here referred as an evaluative audience perception. Because we assume that others will appraise our behavior, we modify our behavior when others are watching as a means to maximize social approval (Allen & Leary, 2007; Engelman, harmann & Tomasello, 2012). While there is ample evidence that we come to develop an evaluative audience perception, questions remain regarding the onset of this phenomenon viewed here as a major cornerstone of human psychology (Rochat, 2009; 2014).

Specifically, when do we begin to perceive others as evaluators of the self and what might contribute to such development?

Recent studies have indicated that behavior modification in the presence of others is present by at least four years of age. Haun and Tomasello (2011) adapted Asch's original conformity experiment, in which individuals publicly express wrong judgments to match that of the majority, to explore conformity in preschoolers. In their study, children would get a coloring book with three different sized tigers on the left and one tiger on the right that varied in size. They were then asked to choose the tiger on the left side that matched the size of the tiger on the right side. Importantly, children were instructed to make this size-judgment decision either publicly by saying the answer aloud, or privately, by pointing silently to what they considered to be the right answer. Mirroring the findings of the original Asch study with adults, children showed a significant tendency to conform to the (wrong) peer majority in the public, but not the private condition. Because the response was directly influenced by whether or not the child's response was known to their peers, the researchers proposed that children conformed to avoid being negatively evaluated by others (Haun & Tomasello, 2011).

More recent studies also indicate that from 5 years children show enhanced performance in the presence of others (i.e., social facilitation, Chevalier et al., 2014). Because fear of evaluation is an implied underlying cause of social facilitation (Cottrell, 1972; Baron-Cohen et al., 2001; Chevalier et al., 2012), Chevalier and colleagues (2014) tested whether social facilitation would be observed in typically developing children as opposed to children diagnosed with autism who display marked decreased attunement to others (Chevalier et al., 2012). Results indicated that typically developing children performed better in theory of mind tasks in the presence of the experimenter, as opposed to being tested alone via a computer. In contrast, the

presence or absence of the experimenter did not affect the performance of children with autism. Importantly, typical and autistic children performed equally in the computer condition. The authors concluded that enhanced performance in the audience condition found in typical children was linked to their sensitivity to others' evaluation of the self. In contrast, children with autism demonstrated a significant lack of social facilitation because they have a diminished concern for another's evaluation (i,e., less of an evaluative audience perception; Chevalier et al., 2012).

The prevalence of an evaluative audience perception in preschool-aged children was further supported in a recent comparative study done by Engelman, Hermann, and Tomasello, (2012). In their study, the researchers manipulated the absence or presence of a peer observer to examine peer influence on five-year olds' propensity to either steal or share. The children were instructed to share stickers with an anonymous peer while they were either alone in the room, or while another peer was watching. Results showed that children shared significantly more and stole fewer stickers when another peer was watching compared to when alone. In contrast, chimpanzees tested in analogous conditions did not show any signs of being influenced by the presence or absence of an experimenter. Taken together, these results confirm that from 4-5 years children begin to display a tendency to adjust their behavior in the presence of others. Importantly, like adults, children tend to behave in a more self-enhancing positive way when being observed by others.

Self-evaluation in Late Infancy

While the evidence in the literature has only recorded an explicit evaluative audience perception in preschool-aged children and beyond, previous research point to emerging signs of self-evaluation by the end of the second year, suggesting that there might be a much earlier onset of an evaluative audience perception (Lewis et al., 1989; Stipek at al., 1992; Rochat Broesch and

Jayne, 2012). Self-evaluation, or children's apparent ability to evaluate their own behavior in positive as opposed to negative terms, emerge from around 18-21 months and are instantiated via self-conscious emotions in front of mirrors (Lewis et al., 1989), or when the child fails to meet a standard or goal (Stipek et al., 1992). Originally studied by Amsterdam (1972), the Mirror Mark Test probes for mirror self-recognition in children, which has become the litmus test for having an objectified and conceptual (represented as opposed to implicit) sense of self (Rochat, 2002; Lewis et al., 1989). In this paradigm, infants are surreptitiously marked with rouge and are then shown a mirror. Consequently, infants (usually around 18 months) who recognize themselves in the mirror will try to remove the rouge from their noses. Importantly, mirror self-recognition is accompanied by a display of self-conscious emotions such as embarrassment, aversion of eye gaze, and blushing (Amsterdam, 1972; Lewis et al., 1989; Lewis, 2001). Such reactions suggest that children passing the Mirror Mark Test are able to evaluate and compare their reflection to their mental representation.

Around the time children begin to pass the Mirror Mark Test, they also begin to manifest a novel sensitivity to norms that are prevalent and well documented by the end of the second year (Kagan, 1981; Rackoszy, Werneken & Tomasello, 2009). Stipek and Colleagues (1992), for example, showed that by 24 months, children self-evaluate against a set standard when presented with a clear fail or success task. In their study, the experimenter would demonstrate how to make different toys work, explicitly delineating a standard of success and failure attempts. After demonstrating both outcomes, the children were free to interact with the toys and encouraged to re-produce success outcomes. Results showed that children evaluated differentially their outcome, displaying embarrassment and gaze aversion when their outcome was a "failure", but smiling and looking at the experimenter when they would succeed.

Because self-conscious emotions such as embarrassment and coyness tend to occur in relation to the presence of other observers (Sattler, 1966; Lewis, 2008; Rochat, 2013), research exploring self-evaluation via the mirror mark task and success/failure tasks in 18-21 month-olds postulate that the manifestation of self-conscious emotions in these tasks could be an outward expression of an emerging sensitivity to others evaluations toward the end of the second year (Stipek et al., 1992; Lewis et al., 1989; Rochat, Broesch & Jayne, 2012; Rochat 2009). However, to our knowledge, this proposal remains a conjecture because studies have yet to directly manipulate and compare infants' responses in relation to the presence or absence of an attentive audience. For example, early manifestations of embarrassment following failure at a task could be related to not meeting the standard or norm, regardless of whether or not another is watching. The question of whether self-evaluation by the second year requires sensitivity to others' evaluation remains open as children might just respond with embarrassment and other self-conscious emotions by simply comparing their own performance to standards and norms they represent independently of others' evaluation.

In an attempt to elucidate on this question, Rochat, Broesch, and Jayne (2012) tested 14-52 month-old infants in one of two conditions: In a *classic Mirror Mark Test condition*, infants were exposed to their specular image after a sticky note was placed on their forehead unbeknownst to them. Alternatively, in a social *norm Mirror Mark Test condition,* another group of same-age infants was tested while the mother and the Experimenter observing the child also wore a sticker on their forehead. Results indicate that the Mirror Mark Test was passed differently depending on the two conditions, whereby infants passing the test in the norm condition were significantly more inclined to touch but leave the mark on their forehead. In contrast, those who passed the Mirror Mark Test in the *classic condition*, tended to touch and

remove the sticky note off their forehead with significantly less hesitation and without any attempts at putting it back on, something often observed in the *norm* condition (Rochat, Broesch, & Jayne, 2012). While signs of conformity in this study suggest that by two years children might be sensitive to others' evaluation, the evidence remains indirect as the child's differential ways of passing the Mirror Mark Test in the classic or norm condition could still be related to the detection of a norm independently of others' evaluation.

Likewise, previous studies (e.g., Lewis et al., (1989); Stipek et al., 1992) do not manipulate systematically the presence or absence of an audience, rendering difficult whether children are actually responding to others' evaluation of the self by this age. The remaining question is still whether self-evaluation in the form of self-conscious emotions before the age of two is in fact an indicator of an early evaluative audience perception. Studies suggest that 4-month-old infants display coyness when others pay insistent attention to them (Reddy, 2004; Reddy, 2000). However, this evidence does not demonstrate evaluative audience perception proper as presence or absence of an audience was not systematically manipulated. As suggested by researcher Vasudevi Reddy: "In predictable contrast with toddlers and adults, 'performance' and potential evaluation was not a relevant context for this expression (coyness) in infancy" (Reddy, 2000, p. 190).

In summary, as of now, the manifestation of self-conscious emotions from around the second year (Lewis et al., 1989; Stipek et al., 1992) only indirectly suggests an early onset of evaluative audience perception. Because of the lack of direct comparison in the context of an audience manipulation, the onset as well as the trajectory of an evaluative audience perception in infancy remains elusive.

The Current Study

In an effort to shed new empirical light on the possible emergence of self-evaluative audience perception, the present study was designed to further explore this possibility that would elucidate on the actual developmental emergence of an evaluative audience perception (4-5 years as it stands, see Englemann, Herrmann & Tomasello, 2012; Haun & Tomasello, 2011), by the time children begin to recognize themselves in mirrors and pass the Mirror Mark Test. We designed two studies with the following general characteristics: In a first study, we created an ambiguous novel toy discovery task, while systematically manipulating the attention of an adult observer toward the infants as they discover and explore the novel toy (a remote controlled robot). In this study, the task was open-ended, with no external norms or standards for the child to follow. The children were simply shown how to activate a toy robot with a remote control.

Our rationale was that if early signs of self-conscious emotions in the second year are indicative of an emerging sensitivity to others' evaluation, we should be able to capture signs of behavior modification as a function of an attentive (Audience Condition) vs. inattentive (No Audience Condition) audience. Inversely, if children display embarrassment because they fail to meet a standard that is independent of others' self-evaluation (i.e., whether they succeed or not in activating the robot), then the presence or absence of an observer should have no modulating effects on the child's behavior. In the latter case, we should not expect any differential emotional expressions or tendency to modulate the frequency of button pressing as a function of attentive vs. inattentive audience condition.

In a second study and as a follow up, we further explored the underlying motivation for infants to modulate their behavior as a function of the presence or absence of an experimenter. If behavior modification in the first study point to infants' emerging sensitivity to others'

evaluation, infants should also manifest strategic signs of behavioral modification as a function of values expressed by the adult during the preliminary demonstration of how the toy robot could be activated via two buttons controlling specific movements. In two alternating demonstrations, one button was associated to a negative outcome with the Experimenter saying, "Oops, oh no!" or a positive outcome with the experimenter saying, "Isn't that great?" We reasoned that an evaluative audience perception as opposed to mere general norm sensitivity should be indexed by the infants' differential enactment of positive vs. negative actions that were modeled by the adult experimenter, depending on whether she is looking or not looking at them (i.e., audience manipulation).

In both studies, children were also tested for their mirror self-recognition using the classic Mirror Mark Test (Amsterdam, 1968; Gallup, 1970; Lewis & Brooks-Gunn, 1979; Bertenthal & Fisher, 1978; Rochat et al., 2012), but testing each child in conditions where the experimenter was either looking or not looking at the child during the test. As a general working hypothesis, assuming that the passing of the Mirror Mark Test (i.e., the child directly touching the mark on his or her face) is an index of self-objectification and explicit self-concept (i.e., the abstract concept of "Me"; Lewis, 1982; Amsterdam, 1968; Rochat, 2003; Rochat & Zahavi, 2011), we predicted that behavioral modification in the Audience condition of the Robot Task would significantly correlate with the passing of the Mirror Mark Test.

As further index of emerging evaluative audience perception, we also predicted that those infants who pass the Mirror Mark Test in both Audience and No Audience Conditions (Study 1, see below), should manifest heightened hesitancy when passing in the Audience Condition measured by the delay to touch the mark from the moment they see themselves in the mirror (see Rochat et al., 2012 for a similar measure). Finally, as another proxy of self-evaluative audience

perception, we also expected heightened signs of embarrassment in the Audience compared to the No Audience Condition in both Robot and Mirror Mark Tests in both studies.

Study 1

Overview: We tested forty-nine 14-24 month-old infants in a novel paradigm (Robot Task) used in parallel to the classic Mirror Mark Test in conditions where the attention of an audience (i.e., the experimenter) was manipulated within-subjects. In the Robot Task, infants were free to play with a novel remote-controlled robot for 30 seconds while the experimenter either observed the child (Audience Condition) or turned 45 degrees sideways and pretended to read a magazine (No Audience Condition). Relative behavioral change in terms of button pressing across conditions was measured. Following this test, infants were tested in two conditions of the Mirror Mark Test with or without the presence of an audience (order counterbalanced). To mark the child in the Mirror Mark Test, a Post-It note was surreptitiously placed on the child's forehead before being confronted with their specular image. Reaching for the mark counted as passing. To examine whether the presence of an observer in the Mirror Mark Test influenced the way in which children would pass the test, latency to pass was recorded in both conditions (i.e., time elapsed from first mirror self-viewing and direct touching of the mark on the body). Furthermore, across conditions and for all tests (Robot and Mirror), we compared the child's emotional response by assessing embarrassment, coyness, happy, fear, and neutral expressions by independent coders on a Likert scale (see below for details).

Method

Participants

A total of seventy-two 14-24 month-old (*M*=18.14, *SD*=2.94; 21 males) healthy infants from predominantly Caucasian middle- to upper middle class family living in the Greater Atlanta

were tested. Participants were recruited through the IRB approved Emory University Child Study Center database. A final sample of 49 participants was included in the analysis. Eleven children were excluded due to experimental error (2); fussiness (6); or parental interference (3). The within-subject design and planned analyses required that all children complete both Audience and No Audience conditions in each task (Robot and Mirror). Because of this requirement, 12 participants were not included since they passed in only one of the audience or no audience condition, not reaching the criterion of either passing or failing in *both* conditions. This criterion was established to allow for the comparison of how infants passed the Mirror Mark Test in the Audience as opposed to the No Audience condition (i.e., delay to pass). Probing our general working hypothesis (see above), this criterion was also meant to allow for more heightened contrast between infants who passed or didn't pass the Mirror Mark Test.

Materials

Robot Task: A red circular remote-controlled toy robot with a 13cm height x 13cm diameter was used for both conditions. The robot was placed in a 32 x 32cm vertical blue box with an opening facing the child, covered and unveiled during test by a yellow cloth. A one button, 13.5cm x 9cm x 2cm rectangular remote control, was placed in a red slip with the activating button exposed. Both the robot and remotes were placed on a 92 x 76cm white table. Each button press made the robot twirl around in place with beeps, sounds, and flashing lights. The activity ended once the button was released.

Mirror Mark Task: A 31 x 31 cm self-standing mirror was placed on the table directly in front of the child for both conditions of the Mirror Mark Test. A yellow cloth initially covered the mirror and was lifted at the beginning of a trial. A 5 x 5 cm yellow 'Post It' note was used and surreptitiously placed on the child's forehead at hairline by the Experimenter affectionately

padding the child's head. Thirty seconds elapsed during which the child showed no sign of noticing the mark prior to uncovering the mirror.

For all subsequent coding, each testing session was video recorded by two cameras, one placed behind the Experimenter providing a frontal view of the child and the other placed behind the child providing a frontal view of the Experimenter.

Procedure

During all of the testing session, the child sat on the mother's lap. Prior to the start of the study, the mother was instructed to remain silent, not intervene, nor to give any hint to her child during test, unless the child showed any signs of distress.

Robot Test: Prior to the actual testing and as a familiarization, the female Experimenter interacted with the child for five minutes, exploring and playing with toys on the ground. After this brief interaction, the child sat on the mother's lap in front of the experimental table (see Figure 3a and 3b for set up). The Experimenter sat 78cm directly across from the child and began the robot demonstration phase by lifting the yellow cloth and uncovering the toy robot on the table. The experimenter then grabbed the remote control and demonstrated how to make the robot work by pressing its button, 3 times in a row, with a 3 second lapse between each press. During demonstration, contingent on the robot moving following each button press, the Experimenter said "Isn't that great?" while looking at the child and smiling. This feedback remained the same following activation of the robot when the child pressed the button. After the demonstration phase, the experimenter conspicuously placed the remote control on the table within reach of the child. In the Audience Condition, the Experimenter looked at the child with a neutral expression. In the No Audience Condition, once the remote was placed in front of the child, she turned 45 degrees sideways and started reading a magazine looking away from the

child. Both conditions lasted 30 seconds in a counterbalanced order across infants with a 1 minute interval. During the interval and at the end of the second condition, the Experimenter got up and said “Okay, we are going to put this away, and I am going to show you something else.”

At the end of the second Robot Task condition, the Experimenter approached the child, looked at the child and said “Can I have the remote?” As the child was looking toward the remote on the table, the Experimenter padded the child’s head, placing surreptitiously the yellow Post-It note on the child’s forehead (see above) in preparation for the Mirror Mark Test, immediately following the Robot Task.

Mirror Mark Test: After the experimenter hid the remote and robot under the table, she placed the self-standing covered mirror on the table facing the child. Once placed on the table, the cloth was removed revealing to children their specular image for 60 seconds. In the Audience Condition, the Experimenter stood behind the mirror facing the child, looking directly at the child from above. In the No Audience Condition, once the mirror was placed on the table in front of the child, the Experimenter conspicuously disappeared behind a large opaque room divider-curtain hanging one meter away to the right of the child. Once out of sight, the Experimenter unveiled the mirror from behind the curtain by pulling a fishing line attached to the covering cloth. The order of Audience vs. No Audience Condition was randomized across participants.

Coding

All coding was done using pre-recorded videos of each task by research assistants who were blind regarding conditions, as well as the working hypotheses. For the Robot Task, frequency of button presses was recorded in real time using JWatcher Event Recorder Software. Presses that successfully activated the robot via both movement and sound, independently of duration, counted as button presses. A second coder coded 20% of participants, with significant

intercoder reliability (correlation; $r = .089$, $p<0.001$).

Three coders coded all participants for emotion using a pre-set criteria based on previous infant and toddler emotion descriptions (Reddy, 2000; Lewis et al., 1989). Across conditions in both Robot and Mirror Tasks, using a likert scale from 0-4 (4 being the highest), the 3 coders rated the child's overall propensity to express respectively embarrassment, coyness, happiness, fear, or neutral emotion. Coders followed the following criteria corresponding to each emotion: 1) *Embarrassment:* presence of gaze aversion, attempts to hide face, and blushing; 2) *Coyness:* leaning back and smiling, or gaze aversion accompanied with smiling; 3) *Happiness*: signs of smiling or giggling; 4) *Fear:* leaning back and widening eyes; 5) Neutral: no distinct emotion identified. Using the same likert scale, Coders also recorded the overall level of engagement (i.e., interest) of the child. Intra-rater reliability for each emotion including engagement received a Chronbach's Alpha value of 0.89 or above.

For the Mirror Mark Test, two coders recorded for all participants whether the child passed or failed the Mirror Mark Test by touching directly the sticky note. Raters agreed on 96% of the whole sample (Kappa=0.928, $p<0.001$). Obliviousness (no touch) of the sticky note counted as not passing for each condition. In addition, the delay to pass in both conditions of the Mirror Mark Test was measured from the time the child made eye contact with the mirror, to the time they first touch the sticky note based on a computer time code with 100^{th} second precision. Raters once again showed significant reliability (Kappa=0.94, $p<0.001$).

Results

Robot Task: In relation to the frequency of button presses in the Robot Task across Audience vs. No Audience conditions, factoring whether the infant passed or failed the subsequent Mirror Mark Test, a 2 (Condition) X 2 (Passing or Failing the Mirror Mark Test)

Mixed-factorial Analysis of Variance yielded a significant Condition-by-Mirror Mark interaction (F (1, 47) =4.956, p=0.03, n^2=0.2). As illustrated in Figure 1a below, pair-wise comparisons indicate that infants passing the Mirror Mark Test displayed significantly greater number of button presses when the Experimenter was inattentive to them in the Robot Test (No Audience Condition), hence showing signs of inhibition when the Experimenter was looking at them (Audience Condition), (F (1, 46) =4.189, n_{p2} = 0.1, p<0.05, see Figure 1a). No audience effect was found when considering infants who failed the Mirror Mark Test. Factoring participants' gender or age did not yield any significant results (respectively F (1, 47) = 0.310, p=0.58 and F (1, 47) =1,413 p=0.241). As discussed below, those children who passed the Mirror Mark Test tended to be significantly older than those who failed. However, the overall inclination to generate button presses on the remote was not different for those children who pass or failed the Mirror Mark Test (F (1, 47) =0.047, p=0.829).

Emotion: Regarding the Robot Task, a repeated measures Analysis of Variance including all five emotions as factors yielded a main effect of Condition for embarrassment only (F (1, 47) =8.574, p=0.005, n^2= 0.2 see Figure 1b), showing that infants displayed significantly more embarrassment in the Audience compared to the No Audience condition, with no significant interaction nor any main effects of age, gender, or more importantly, whether infants passed or failed the Mirror Mark Test (all p>0.05). Likewise, an overall Condition main effect in relation to embarrassment was found in the subsequent Mirror Mark Test (F (1, 47) =5.164, p=0.028, n^2=0.2).

Mirror Mark Test: When considering all participants tested in the Mirror Mark Test in both conditions (N=61, see participants above), 24 infants passed in both Audience and No Audience conditions, as 12 passed in only one of the two Conditions. Noteworthy is the fact that

out of the 12 participants passing the test in only one Condition, 10 passed it in the No Audience condition (Binomial Test $p<.02$).

For those participants who passed in both conditions (N=24), we compared how they passed the test as a function of condition by measuring the delay to touch the mark on their forehead from the moment the mirror was unveiled and the infant saw their own reflection. Again, this index was used as a proxy of relative hesitation to touch the mark, expected to be significantly longer in the Audience compared to the No Audience Condition, following our working hypothesis (see also Rochat et al., 2011 for comparable analysis and findings).

Paired-sample t-test comparing latency (in seconds) to pass between Audience and No Audience conditions shows that infants who passed the Mirror Mark Test in both conditions (N= 24) showed significantly more hesitation (delay or time lag) before touching the mark when the Experimenter was attentive as opposed to inattentive toward the infant (t (23) =2.318, p=0.03, d=0.53, see Figure 1c). Controlling for the total amount of time spent looking in the mirror yielded no significant result (t (23) =1.644, p=0.114). In addition, as already mentioned, age (β = 0.304, t (48) = 3.098, p= 0.003) but not gender (x^2 (1, N= 24) =3.6, p=0.085) was a predictor of passing the Mirror Mark Test, with older infants showing significantly more passing. When factoring the delay to pass in the Mirror Mark Test in the overall Analysis of Variance on button presses in the Robot Task, this variable yielded no significant results (i.e., did not predict button pressing as a function of condition).

Discussion

By systematically manipulating the attention of the experimenter toward participants (Audience vs. No Audience condition) in a within-subject design, results of this first study provide compelling evidence that by the end of the second year, infants begin to display an

evaluative audience perception. Specifically, we found that infants systematically modifying their behavior (button presses, relative expression of embarrassment, delay to pass the Mirror Mark Test) as a function of others' attention. Furthermore, passing the Mirror Mark Test predicted the audience condition effect in the Robot Task, albeit with a significant age confound.

While these results confirm our working hypotheses on the emergence of an evaluative audience perception by the end of the second year, three confounds need to be considered. The first is related to variations in infants' temperament, in particular the relative shyness of the participant. Behavioral inhibition in the presence of others is indeed the trademark characteristic of an infant's shy temperament (Rothbart & Mauro, 1990). While we think that it is unlikely that all children who passed the Mirror Mark Test (which was positively correlated with behavior modification between conditions), also happened to be those with a particularly shy temperament, questions remain on whether temperament could predict relative inhibition across Audience vs. No Audience condition in the Robot Task. Interestingly, 12 children who were excluded from the analysis of the first study because they failed to pass in both Mirror Mark Test conditions also showed overall reduced button pressing independent of conditions and when compared to those children included in the final sample for our analysis. It is therefore possible that inter-individual differences of temperament play a role in our findings. In the second study, we factored each infant temperament as assessed by the short temperament assessment form of the Early Childhood Questionnaire (Putnam, Gartstein, and Rothbart, 2006).

A second possible confound is that infants did show inhibition in the Audience Condition because the testing situation was too ambiguous, the Experimenter providing no particular instructions after placing the remote on the table. It is indeed possible that in the audience condition, such ambiguity might be simply heightened. One could argue that rather than it being

an evaluative audience perception, the infant simply manifest inhibitory response linked to the open ended testing situation. In Study 2, we address this question by making the robot situation less ambiguous and give more instructions to infants following the demonstration. As she hands the remotes to the child, the Experimenter does instruct the child, "Your turn!"

Importantly, we added more structure to the task by demonstrating two different outcomes of robot activity with two different remotes, one positively reinforced by the Experimenter during demonstration, the other negatively reinforced (see method below). In addition to alleviating ambiguity, this two-button choice allowed for further probing of an emerging evaluative audience perception in the Audience vs. No Audience condition. In addition, we added a control group where no differential value feedback was given during demonstration by the Experimenter. The rationale was that, controlling for age, evaluative audience perception would be further demonstrated if, compared to the Control group, the Experimental group shows signs of being significantly more discriminant and button pressing as a function of Audience vs. No Audience condition.

Finally, in our sample, infants passing the test were significantly older than those who failed the Mirror Mark Test. Therefore, while we did find a significant link between infants who passed the test and their extent to which they modify their behavior as a function of Audience vs. No Audience condition, it is possible that a major factor is age as a global variable. Without controlling for age, it is not possible to demonstrate any specific link between the manifestation of a conceptual self and an emerging evaluative audience perception. In Study 2, based on a large sample of tested participants (N=62), we address this issue by systematically matching age and gender of those participants that either passed or failed the Mirror Mark Test in both the Experimental and Control (see participants below). Furthermore, for Study 2, we used a less

stringent criterion for determining who passed the Mirror Mark Test (passing in at least in one of the two conditions), in order to keep reasonable sample sizes and not lose power in our effort to match age and gender within each group. As a consequence, we were unable to do any within subject comparison of conditions in relation to the Mirror Mark Test.

Study 2

Participants

A total of 51 participants were included, with 31 participants in the Experimental group (M=20.64, SD=2.83; 10 males) and 20 participants in the Control group (no value feedback regarding the two remotes, see below; M= 19.9, SD=3.3; 12 males). There were no significant differences in age, gender, or temperament between control and experimental group (p>0.05). Four additional children were tested but were excluded due to fussiness.

Material

The material was the same as Study 1, except for two remote control devices used in the Robot Task as well as a change in the kind of mark used in the Mirror Mark Test (odorless rouge smeared on the infant's nose). Post-It mark was changed for rouge because the Post-It often did not adhere on certain infants' skin, notwithstanding the fact that smeared rouge is what has been typically used in other studies (Amsterdam, 1972; Lewis & Brooks-Gunn, 1979; Bertenthal & Fisher, 1978).

Procedure

Robot Task: The basic procedure was identical to Study 1, except that a pair of distinct remote control devices was placed side-by-side in a clear tray on the table. Each remotes' button activated different light flashing and sound-producing motions in the robot. For the Experimental

group, during demonstration, each remote was associated by either a positive "Isn't that great?" or negative verbal comment ("Oh oh! Oooops! Oh no!") by the Experimenter.

As in Study 1, the robot was initially placed in an open vertical box facing the child. In 2 alternated demonstrations of 3 seconds, one remote spun the robot in circle inside the box. In contrast, the second remote made the robot move forward and step out of the box, with analogous sounds and flashes. Positive or Negative verbal comments by the Experimenter during demonstration for either remote was counterbalanced across participants.

For the Control group, the Experimenter demonstrated each remote in alternate succession as for the Experimental group, but with no distinguished verbal comments associated with either remote. Following the action of either remote, the Experimenter said, "Oh wow!" The color of remote, type of verbal feedback, alternating order of remote demonstration, as well as left/right location of the remote on the tray were counterbalanced across infants of each group. For both groups, in contrast to Study 1, immediately following the last demonstration and pushing the tray containing the pair of remotes toward the infant, the Experimenter looked conspicuously toward the infant saying "Your turn!" The experimenter then turned 45 degrees and pretended to read a magazine for 30 seconds (No Audience Condition) or looked at the child with a neutral face for 30 seconds (Audience Condition). At odd with the procedure used in Study 1, after the demonstration and during the test phase in either condition, the Experimenter refrained from giving any further verbal feedback to the infants following each button press.

Mirror Mark Test: Immediately following the Robot Task, each infant was tested in the Mirror Mark Test in both Audience and No Audience Conditions, with order counterbalanced across infants of each group. Odorless rouge was surreptitiously smeared on the child's nose at the beginning of testing. In the Audience Condition, the Experimenter kneeled in front of the

mirror behind the Child, reflecting her own image in the mirror while staring at the infant's reflection (Audience Condition), or turning her back to the specular image (No Audience Condition).

Coding

The same dependent measures as in Study 1 were used, adding the button presses for either the positive or negative outcome remotes in the Robot Test. Reliability testing on 20% of all participants yielded high coefficient correlation for all emotion coding (> or = 0.946); passing of the Mirror Mark Test (K=0.844, r=0.86, p<0.001); as well as button presses for either remote (*K*=0.989, *p*<0.01) in the Robot Task.

Results

Robot Task: To measure the amount of button presses from each remote between conditions, we conducted a 2 (Condition: Audience vs. No Audience) X 2 (Remote: Positive vs. Negative) factorial Analysis of Variance and entered two variables as between subject factors: group (Experimental vs. Control) and Mirror Mark Test (Passing vs. No Passing). Results yielded a significant three-way interaction of group, condition, and remote (*F* (1, 49) =4.394, *p*=0.041, n^2=0.1, see Figure 2a), with no significant interaction of whether infants passed the Mirror Mark Test (p>.05). Factoring age in the overall analysis of variance on button presses did not yield any significant effect (*F* (1, 47) =.043, *p*=.836).

A follow-up paired- samples t-test indicated that in the Control group, there were no significant differences in button pressing for the blue remote (*t* (19) =0.5, *p*=0.623) nor the orange remote (*t* (19) =1.116, *p*=0.278) between conditions. Importantly, in the Experimental Group, children pressed the positive remote significantly more in the Audience condition (*t* (30) =2.177, *p*= 0.03), but pressed the negative remote more in the No Audience condition (*t* (30) =

1.988, p=0.04, see Figure 2a).

In all, these results indicate that regardless of age, only children in the Experimental Group demonstrate signs of an evaluative audience perception by selectively preferring to press either remote as a function of Condition. When no value was given to either remote by the Experimenter during demonstration (Control Group), this selectivity disappears. Furthermore, for the Experimental Group, infants appear to shift systematically in preferring to press the positive remote in the Audience Condition and the negative remote in the No Audience Condition, strongly supporting the idea that, as hypothesized, infants by their second birthday are sensitive to other's evaluations.

Temperament: A correlation matrix on the relation between total button pressing and the distinct dimensions (Sociability, Inhibitory Control, Fear, and Shyness) of the short parental temperament questionnaire by Putnam, Gartstein, and Rothbart's (2006) yielded a significant relation between shyness and total button presses for both conditions (r =-0.467, p=0.029). However, follow-up analyses showed that Shyness was not a significant covariate when factored in the larger Analysis of Variance (F (1, 29) =0.464, p=0.503).

Emotion: When factoring all 5 measures of emotion, including embarrassment, the analysis did not yield any significant main effect, nor any interaction in either the Robot Task or the Mirror Mark Test (all p>0.05).

Discussion

In the second study, for the Robot Task, we explored whether infants would be strategic in their button pressing as a function of audience when the Experimenter first demonstrated a positive versus negative value for actions coming from each remote. We predicted that if infants were truly modifying their behavior in Study 1 as an expression of an evaluative audience

perception, then we should also find evidence that infants express differential inclination in activating the positive vs. negative outcome. Confirming our hypothesis, results show that infants of the Experimental Group pressed the positive remote significantly more in the Audience compared to the No Audience Condition. Inversely, they tended to press significantly more the negative remote in the No Audience Condition. For the Control Group, who had no preliminary verbal feedback regarding the value associated with the remotes, infants did not show any significant preference, regardless of condition. Following our rationale and hypotheses, infants of the Experimental Group appeared to be indeed strategic in their choice of remote, depending on whether the Experimenter was attentive or inattentive to them. This finding further demonstrates the existence of an evaluative audience perception by 24 months. In addition, we confirm that the relative shyness temperament of the infant as reported by parents did not predict infants' relative proclivity to activate the robot in general. Therefore, in the context of our study, we did not find that the degree of an infant's shyness as reported by parents is a significant predictor of an evaluative audience perception in infancy.

Contrary to Study 1 where age was not controlled, results of this second study indicate that the passing of the Mirror Mark Test is not predictive of differential button pressing as a function of either group (Experimental vs. Control), or condition (Audience vs. No Audience). Therefore, when controlling for age, and contrary to what we hypothesized, the passing of the Mirror Mark Test (presumably a measure of self-concept or self-objectification) is not linked, nor does it predict what we construe here as evidence of an evaluative audience perception. Note, however, that the criterion for passing the test was less stringent in Study 2 (passing at least in one of the two conditions), compared to Study 1 (passing in both conditions). This might have contributed to this absence of significant correlation. Future studies with varying criteria would

help elucidate further on the actual link between self-concept as measured by the Mirror Mark Test and emerging self-evaluative audience perception. Other tests of self-concept might be needed, such as the body as an obstacle test that has been used with same age infants by Moore et al. (2007).

Finally, contrary to what we found in Study 1, the fact that there was no significant difference in our coding of embarrassment between Audience vs. No Audience Conditions in both the Robot Task and the Mirror Mark Test could rest on the fact that for Study 2, a direct instruction was given to the infants in the Robot Task ("Your turn!") as the remotes were pushed toward them. We propose that his simple instruction lifted the ambiguity of the test, giving more meaning and less reasons for the kind of behavioral inhibition or freezing that is typically associated with embarrassment.

General Discussion

It has been widely reported that by the preschool year, children begin to strategically modify their behavior in the presence of others as means to maximize self-presentation (Engelman, Herrman, & Tomasello, 2014; Chevalier et al., 2014). Importantly, this continues throughout development (Banarje, Bennett, & Luke, 2012) and into adulthood (Walting & Banerjee, 2007; Schlenker & Leary, 1982). While our proclivity to modify our behavior is ubiquitous, the onset is underspecified. This study thus explored when infants would begin to systematically modify their behavior in the presence of an attentive audience, an index of an evaluative audience perception. By directly manipulating the attention of an audience in a novel paradigm, we found that the first signs of an evaluative audience perception emerge by 24 months, a much earlier onset than previously found (Engelmann, Hermann, & Tomasello, 2014).

In Study 1, we found that children would display inhibition to play with a novel toy when the experimenter was watching, even when there was no clear norm or standard that prevented them from doing so. In addition, we also found that children displayed more embarrassment when the experimenter was attentive than when she was inattentive. The results of the second study further explored why children modified their behavior in the presence of an observer in Study 1 by exploring whether children would strategically choose to activate a button that was associated with more positive feedback as a function of audience. Confirming our initial hypothesis, children chose to press the positive button significantly more when the experimenter was watching and chose to press the negative remote when the experimenter was not watching. We consider this strong evidence supporting that infants begin to be sensitive to others' evaluation by the end of the second year.

The emergence of an evaluative audience perception marks a major qualitatively change in how the attention of others is perceived in development. It has indeed been widely documented that children are influenced by another's attention from two months of age. For example, children will display distress when another disrupts normal interaction and poses a still-face (Tronick et al. 1972). By four months, infants will also display coyness or distress when another fails to interrupt direct contact for an extended period of time (Reddy, 2000; 2004). Therefore, the attention of others does influence the emotion of the infant, but not in the context of an evaluative audience. In the first year, children are simply responding to expectations they have of *others*, not of whether or not they themselves meet the expectations of others. It is not until the second year that trademarks of an evaluative audience perception, such as self-conscious emotions, self-evaluation, and now behavior modification, begin to emerge (Stipek et al., 1992; Lewis et al., 1989; Kagan, 1981).

Additionally, by nine months, infants also show the capability of considering another's evaluation about an *external* object or circumstance via social referencing (Campos and Stenberg, 1981; Rochat & Striano, 2000). The original experiment by Campos and Sternberg (1981) showed that 9-12 month old infants would look to the caregiver before crossing a visual cliff to get a toy from the other side. Importantly, infants incorporated the caregiver's facial feedback, and would cross the visual cliff significantly more when the mother was smiling compared to when the mother displayed fear (Campos & Sternberg, 1981). Striano and Rochat (2000) also demonstrated how children take into account the attention of an experimenter in relation to social referencing. The researchers found that 10-month- old infant's social referencing significantly decreased when the experimenter was turned sideways as opposed to when the experimenter was facing the child. Crucially however, the fact that the experimenter was turned sideways did not modify the way in which children interacted with the novel toy. It simply reduced the amount of times the child would refer to the experimenter, because they understood experimenter could not see them or the toy in the inattentive condition. Likewise, infant's behavior in social referencing paradigms is not modified based on an assumption that their behavior will evoke a positive or negative appraisal from a caregiver. Instead, their behavior is influenced by feedback provided about *external* circumstances (i.e., is this safe or not safe?). This is very different from changing one's behavior because someone is observing, which is what occurred in both of our studies.

While our two studies grant support that an evaluative audience perception emerges toward the end of the second year, several questions remain in regards to children's sensitivity to other's potential evaluations. For example, what is the mechanism underlying self-consciousness? Contrary to our initial hypothesis as well as assumptions from previous work

(Stipek at al., 1992; Lewis et al., 1989), mirror-self recognition does not seem to be a necessary precursor to develop an evaluative audience perception or self-conscious emotions. Although behavior modification in Study 1 was highly correlated with mirror self-recognition, this effect dissipated in our second study when we controlled for age. Therefore, it is possible that a separate mechanism gives rise to novel sociocognitive changes emerging toward the end of the second year.

Another question is, how and when do we begin to integrate another's feedback of our behavior into our self-concept? Our perception of how others view us becomes an integral part of our own self-view, a widely supported concept known as the "looking-glass self" (Cooley, 1902; Mead, 1934). In regards to our self-concept, philosopher Charles Horton Cooley insightfully stated, "I am not who you think I am; I am not who I think I am. I am who I think you think I am" (Cooley, 1902, p 202). While developing an evaluative audience perception could be considered as a first step to being able to develop a holistic self-view, it is unclear when and how we begin to incorporate other's perception of us into our own self view. Much work remains to be done to fully comprehend the impact of self-evaluative audience perception on self-concept, beyond its first clear manifestation by the end of infancy.

Conclusion

Empirical and anecdotal evidence shows that humans become to perceive others as evaluators of the self. This perception often motivates us to modify our behavior when another is watching as a ways to maximize self-presentation. While this phenomenon has been recorded in four-year old children as well as adults, the developmental origins of an evaluative audience perception had yet to be explored. To address this gap in the literature we tested 14-24 month-olds, an age in which self-conscious emotions emerge, to see whether behavior modification

would also be observed at this age. Our results confirmed that children who passed the Mirror Mark Test also demonstrated behavior modification, specifically inhibition, when there was an observer. Additionally, our second study showed that children are strategic in the way they behave when there is an audience, choosing the button that elicited positive feedback from the experimenter significantly more when she was looking as opposed to when she was not looking.

In light of these findings, there are critical methodological and theoretical implications. If by the end of the second year infants do have an evaluative audience perception (i.e., other's presence will influence their behavior) then it calls into question the common practice of having several adults present during studies with infants, and whether this induces bias in these measures. More consideration should be taken in relation to how the presence of others may affect the child's emotion and behavior in a task. Of theoretical importance, studying the typical onset and developmental trajectory of an evaluative audience perception could inform our understanding of clinical disorders characterized by diminished or heightened sensitivity to evaluation such as autism or social anxiety (Shlenker & Leary, 1982; Chevalier, 2012). By investigating when, how, and what factors contribute to our sensitivity to other's evaluation, better interventions could potentially be implemented. While the two studies we address here suggest the ontogeny of human sensitivity to other's potential evaluations, much research remains to fully explore the trajectory of this defining human characteristic.

References

Amsterdam, B.K. (1972). Mirror self-image reactions before age two. *Developmental Psychobiology*, 5 (1972), pp. 297–305

Amsterdam, B. K. (1968) Mirror Behavior in Children Under Two Years of Age Doctoral Dissertation, Univ. North Carolina. Order No. 6901569; University Microfilms, Ann Arbor, MI. 48106.

Asch, S. E. (1956). Studies of independence and conformities. *Psychological Monographs*, 70, No. 9 (Whole No. 416).

Banerjee, R., Bennett, M., & Luke, N. (2012). Children's Reasoning About Self-Presentation Following Rule Violations: The Role of Self-Focused Attention. *Child Development*, 83(5), 1805-1821. doi:10.1111/j.1467-8624.2012.01813.

Bertenthal, B., & Fisher, K. (1978). Development of self-recognition in the infant. *Developmental Psychology*, 14, 44-50.

Chartrand T.L., Bargh J.A. (1999). The Chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, 76, 893–910

Chevallier, C., Parish-Morris, J., Tonge, N., Le, L., Miller, J., & Schultz, R. T. (2014). Susceptibility to the audience effect explains performance gap between children with and without autism in a theory of mind task. *Journal Of Experimental Psychology: General*, *143*(3), 972-979. doi:10.1037/a0035483

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, 16(4), 231-239.

Cooley, C. H. (1992). *Human nature and the social order*. Transaction Publishers.

Cottrell, N. B., Wack, D. L., Sekerak, G. J., & Rittle, R. H. (1968). Social facilitation of

dominant responses by the presence of an audience and the mere presence of others. *Journal of Personality and Social Psychology*, 9(3), 245.

Dana, J., Weber, R., & Kuang, J. (2007). Exploiting moral wiggle room: experiments Demonstrating an illusory preference for fairness. *Economic Theory*, 33(1), 67-80. doi: 10.1007/s00199- 006-0153-z

Engelmann, J. M., Herrmann, E., & Tomasello, M. (2012). Five-year olds, but not chimpanzees, attempt to manage their reputations. PloS one, 7(10), e48433.

Gallup, G. G. (1970). Chimpanzees: Self-recognition. *Science*, 167, 86-87.

Goupil, L., Romand-Monnier, M., Koudier, S., (2016). Infants ask for help when they don't know. *PNAS, 113 (13).* doi*:* 10.1073/pnas.1515129113

Haun, D., & Tomasello, M. (2011). Conformity to peer pressure in preschool children. *Child Development*, 82(6), 1759-1767.

Kagan, J. (1981). The second year: The emergence of self-awareness.

Leary, M., & Allen, A. (2006) Social Motivation (pp. 41-59). New York: Taylor and Francis Group.

Lewis, M., & Brooks-Gunn, J. (1979). *Social cognition and the acquisition of self*. New York: Plenum Press.

Lewis, M., Sullivan, M. W., Stanger, C., & Weiss, M. (1989). Self-development and self-conscious emotions. *Child development*, 146-156.

Mead, G. H. (1934). *Mind, self, and society.* Chicago: University of Chicago Press.

Moore, C., Moore, C., Mealiea, J., Garon, N., & Povinelli, D. J. (2007). The development of body self-awareness. *Infancy*, *11*(2), 157-174.

Putnam, S. P., Gartstein, M. A., & Rothbart, M. K. (2006). Measurement of fine-grained aspects of toddler temperament: The Early Childhood Behavior Questionnaire. *Infant Behavior and Development,* 29 (3) , 386-401.

Rakoczy, H., Warneken, F., & Tomasello, M. (2009). Young children's selective learning of rule games from reliable and unreliable models. *Cognitive Development, 24*(1), 61-69.

Reddy, V. (2000). Coyness in early infancy. *Developmental Science*, *3*(2), 186-192.

Rochat, P., & Striano, T. (2000). Perceived self in infancy. *Infant Behavior and Development*, 23(3), 513-530.

Rochat, P. (2003). Five levels of self-awareness as they unfold early in life. *Consciousness and Cognition, 12,* 717-731

Rochat, P. (2009) Others in Mind – Social origins of self-consciousness. N.Y. Cambridge University Press.

Rochat, P., & Zahavi, D. (2011). The uncanny mirror: A re-framing of mirror self-experience. *Consciousness and Cognition*, 20(2), 204-213.

Rochat, P., Broesch, T., & Jayne, K. (2012). Social awareness and early self- recognition. *Consciousness and Cognition*, 21(3), 1491-1497.

Rochat, P. (2013). The gaze of others. In M. Banaji & S. Gelman (Eds.), Navigating the SocialWorld: What Infants. Children, and Other Species Can Teach Us (205-211). New York, NY: Oxford University Press

Rothbart, M. K., & Mauro, J. A. (1990). Temperament, behavioral inhibition, and shyness in childhood. In *Handbook of social and evaluation anxiety* (pp. 139-160). Springer US.

Sattler, J. M. (1966). Embarrassment and blushing: A theoretical review. *Journal of Social Psychology, 69,* 117-133

Schlenker, B. R., & Leary, M. R. (1982). Social anxiety and self-presentation: A conceptualization model. *Psychological bulletin*, *92*(3), 641.

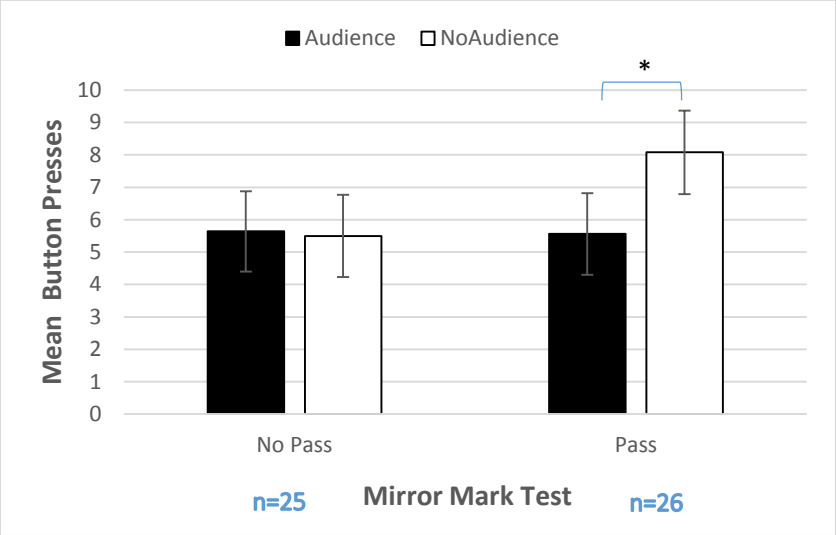
Stipek, D., Recchia, S., McClintic, S., & Lewis, M. (1992). Self-evaluation in young children. *Monographs of the society for research in child development*, i-95.

Triplett, N. (1898). The dynamogenic factors in pacemaking and competition. *The American Journal of psychology*, *9*(4), 507-533.

Watling, D., & Banerjee, R. (2007). Children's understanding of modesty in front of peer and adult audiences. *Infant and Child Development*, *16*(3), 227-236.

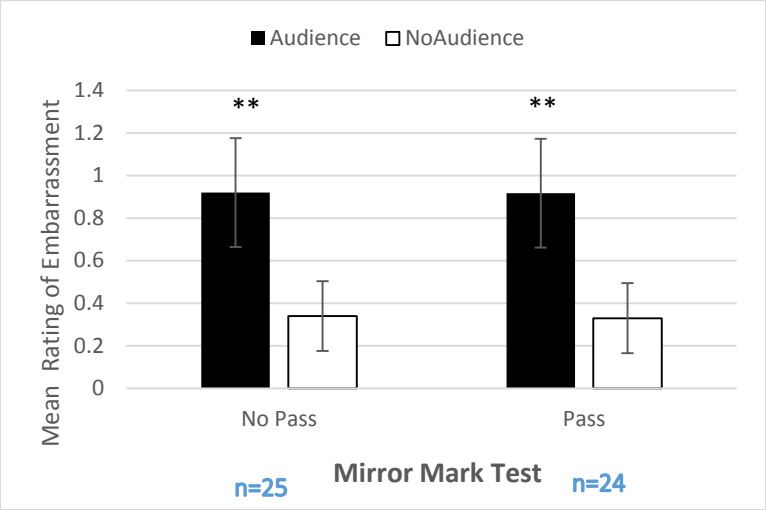
Zajonc, R. B. (1965). *Social facilitation*. Research Center for Group Dynamics, Institute for Social Research, University of Michigan.

Figure 1a: Average Button Pressing in Audience vs. No Audience Condition as a Function of Passing the Mirror Mark Test



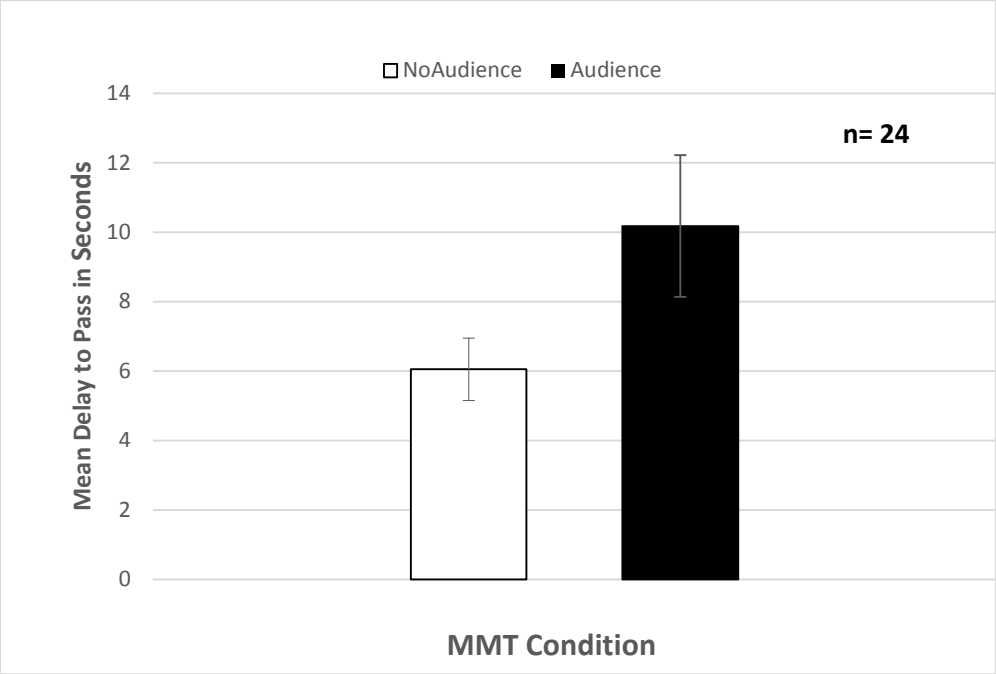
Note: Mixed factorial ANOVA yielded a significant interaction between passing of the Mirror Mark Test and condition of the Robot Task, where infants who passed the Mirror Mark Test pressed the button significantly more when the experimenter was inattentive (*M*= 8.08) versus attentive (*M*=5.39). Error bars represent standard error. *=p< 0.05

Figure 1b: Display of Embarrassment in Robot Task by Condition



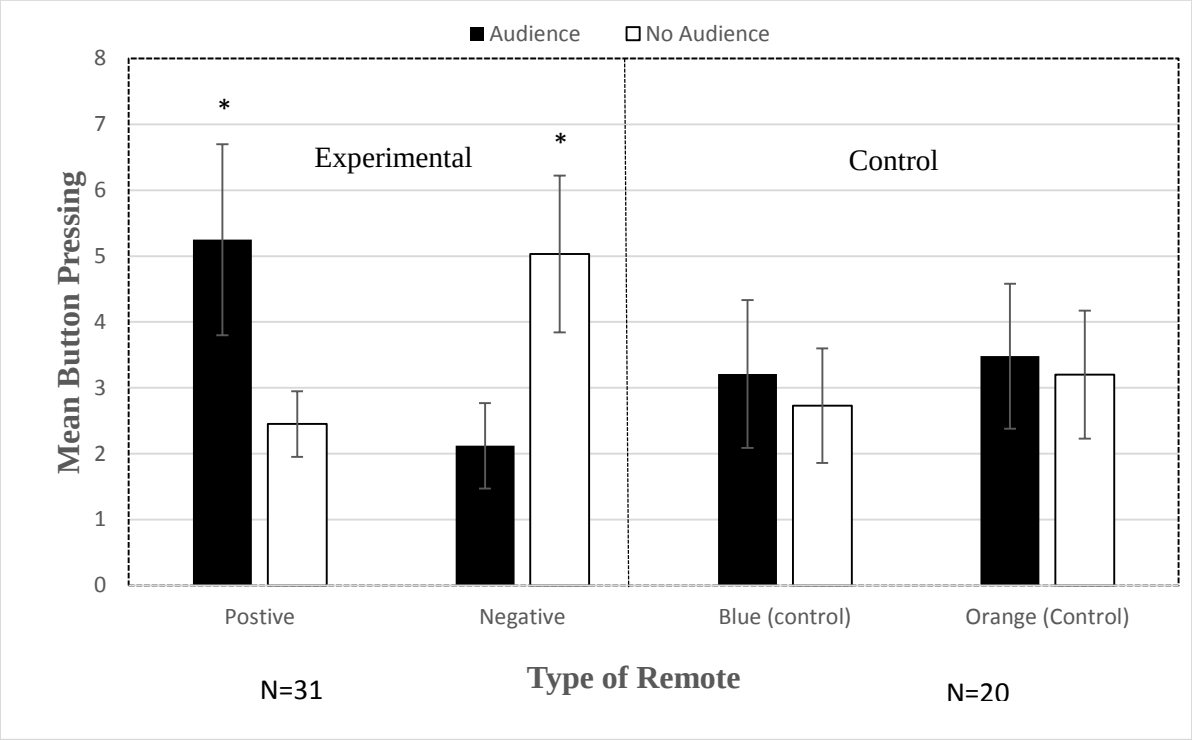
Note: Overall, infants display significantly more embarrassment in the audience (i.e, experimenter is attentive) versus the no audience condition. Error bars represent standard error. **=p< 0.01)

Figure 1c: Delay to Pass the Mirror Mark Test in Audience vs No Audience Conditions



Note: The delay to touch the mark in the Mirror Mark Test for those who passed was recorded in seconds. Analysis indicated that children took significantly longer to pass in the audience condition, further indicating inhibition.

Figure 2a: Average Button pressing as a Function of Condition between Experimental (Positive/Negative) and Control (Blue/Orange) Group



Note: Overall, children in the Experimental group pressed the positive button significantly more in the Audience condition, but pressed the negative button significantly more in the No Audience Condition. In contrast, children in the control group did not prefer one remote over the other. Error bars represent standard error. *=p< 0.05

Figure 3a: Robot Task Conditions

Robot Demonstration	No Audience Condition	Audience Condition

Figure 3b: Bird's Eye View of Room Set-Up