

## **The Effects of the Establishment of Adult Faces and/or Voices as Conditioned Reinforcers for Children with ASD and Related Disorders**

Jacqueline Maffei\*, Jessica Singer-Dudek<sup>1</sup> & Keohane Dolleen-Day \*\*

\*Columbia University Graduate School of Arts and Sciences, \*\*Nicholls State University

### **Abstract**

We tested the effects of the establishment of conditioned reinforcement for observing human faces and/or voices on the rate of learning, observing responses, and verbal operant emissions for four children, ages 4-5 years, with autism spectrum disorders (ASD) and related disorders. We used a non-concurrent, delayed probe design across participants with pre and post-intervention measures. The intervention included a conjugate stimulus-stimulus pairing procedure. Results demonstrated that as a function of the intervention, faces were conditioned for three out of three participants and voices were conditioned for two out of two participants for whom either was lacking respectively prior to the intervention (both faces and voices were conditioned for one participant). Post-intervention probes demonstrated increases in rate of learning, observing responses, and verbal operants for all four participants.

**Keywords:** Verbal Developmental Cusps, Conjugate Reinforcement, Stimulus-stimulus Pairings, Conditioned Reinforcement, Faces, Voices.

## **El Efecto del Establecimiento de Caras y/o Voces de Adultos como Reforzadores Condicionados para Niños con TEA y Desórdenes Relacionados**

### **Resumen**

Se probaron los efectos del establecimiento del reforzamiento condicionado al observar caras o voces de humanos sobre la tasa de aprendizaje, de la tasa de respuestas de observación y sobre la emisión de operantes verbales de cuatro niños de 4 a 5 años con trastorno del espectro autista (ASD, por sus siglas en inglés) y de trastornos relacionados. Se utilizó un diseño no concurrente demorado entre participantes con mediciones pre y post intervención. La intervención incluyó un procedimiento de apareamiento conjugado estímulo-estímulo. Los resultados mostraron que, como función de la intervención, las caras se condicionaron para tres de tres participantes y las voces se condicionaron para dos de dos participantes, para quienes dicho condicionamiento no estaba presente antes de la intervención (ambas, las caras y las voces ya estaban condicionadas para uno de los participantes). Los sondeos post-intervención demostraron un incremento en la tasa de aprendizaje, de las respuestas de observación y de las operantes verbales para los cuatro participantes.

**Keywords:** Cúspides de Desarrollo Verbales, Reforzamiento Conjunto, Apareamientos Estímulo-Estimulo, Reforzamiento Condicionado, Caras, Voces.

Original recibido / Original received: 15/07/2014

Aceptado / Accepted: 25/09/2014

---

<sup>1</sup> This study was conducted as part of the first author's doctoral dissertation under the mentorship of the second and third authors.

Correspondence: Programs in Teaching as Applied Behavior Analysis, Teachers College, Columbia University, Box 22, 525 West 120th Street, New York, NY 10027, E-mail: js408@tc.columbia.edu

Observing the human eyes, or face, is one of the most crucial early observing responses and what many propose is the first step to becoming verbal (Arnold, Semple, Beale, & Fletcher-Flinn, 2000; Baron-Cohen, Baldwin, & Crowson, 1997; Cleveland, Kobiella, & Striano, 2006; Kleinke, 1986). Others agree that eye contact is important for infant-adult interaction, socialization, and cognitive development (Senju, Kikuchi, Hasegawa, Tojo, & Osanai, 2008; Symons, Hains, & Muir, 1998). While most research has focused on eye gaze or eye contact, other studies have shown that the movement of the lips, jaw, face, and tongue also select out observing responses and aid in communication and learning (Kleinke, 1986; Massaro & Bosseler, 2006; Mirenda, Donnellan, & Yoder, 1983; Striano & Bertin, 2004). Neuro-typical infants 7 to 11 weeks old were found to scan the eye area of faces more intently when voices were introduced (Haith, Bergman, & Moore, 1977). For children with autism, this critical observing response is often missing (Baron-Cohen et al., 1997; Ellsworth, Muir, & Hains, 1993; Hains & Muir, 1996; Senju et al., 2008). Recent research indicates that deficits in eye contact can be detected in infants as young as 2 to 6 months of age and may be indicative of a later diagnosis of autism (Jones & Klin, 2013).

Researchers focusing on verbal behavior development have identified several components that appear foundational to becoming verbal. These include conditioned reinforcement for observing adult faces, listening to adult voices, and observing two- and three-dimensional stimuli in the environment (Greer, Pistoljevic, Cahill, & Du, 2011; Keohane, Luke, & Greer, 2008; Keohane, Pereira Delgado, & Greer, 2009; Pereira Delgado, Greer, Speckman, & Goswami, 2009). Observing people and objects in the environment provides a context for individuals to participate in verbal exchanges with one another. These observing responses are operants, selected out by their reinforcers; thus, the stimuli that are observed must be conditioned reinforcers (Dinsmoor, 1983). Therefore, it is the establishment of the reinforcer for observing that is the critical foundation for verbal development.

When observation of the human face is missing from an individual's community of reinforcers, the individual will likely not respond to or even observe the presence of another individual, let alone verbal antecedents delivered by another (speaker). A child who lacks conditioned reinforcement for human faces and/or voices does not orient toward others, whether they are speaking or not speaking, and is not likely to respond to greetings or instructions from a speaker. A child at this level of verbal capability will most likely present at a pre-listener level of verbal behavior (Greer, 2002; Greer & Keohane, 2005; Greer & Ross, 2008; Skinner, 1957).

When individuals lack observing responses for human faces and/or voices they are not able to contact reinforcement from the presence of other individuals, as a speaker or a listener, and subsequently opportunities for contacting other social contingencies are limited. Conditioned reinforcement for observing faces and voices is foundational to increases in the complexity of verbal development and when it is missing further verbal development is not possible.

The observation of human faces and other observing responses meet the definition of what Rosales-Ruiz and Baer (1997) call behavior developmental cusps. These behavior developmental cusps, which include observing responses

as well as behaviors such as crawling and walking, are important developmental stages that, once attained, allow children to progress in ways they could not prior to their attainment. Once established, the individual is afforded opportunities to contact new environmental contingencies, and new reinforcers, and as a result learn new skills that support the performance of more complex tasks (Greer & Keohane, 2005; Greer & Ross, 2008; Greer & Speckman, 2009). Therefore, children can learn things they could not before because they can contact new contingencies. Children can also learn new things faster due to accelerated establishment of stimulus-response relations, or stimulus control. In summary, verbal cusps allow children to contact new conditioned reinforcers that subsequently lead to accelerated rates of learning; identifying and inducing missing verbal developmental cusps is crucial for the development of complex levels of verbal behavior.

The establishment of new conditioned reinforcers often takes place via stimulus-stimulus pairings. The stimulus-stimulus pairing procedure has been used to expand children's community of reinforcers by conditioning non-preferred stimuli as reinforcers, resulting in new responses. Such responses include looking at books, playing with toys, observing two-dimensional stimuli, responding to human voices, and the emission of new vocal sounds (Greer, Becker, Saxe, & Mirabella, 1985; Greer, Dorow, Wachhaus, & White, 1973; Longano & Greer, 2006; Miguel, Carr, & Michael, 2002; Nuzzolo-Gomez, Leonard, Ortiz, Rivera, & Greer, 2002; Pereira Delgado et al., 2009; Rheingold, Gerwirtz, & Ross, 1959; Smith, Michael, & Sundberg, 1996; Sundberg, Michael, Partington, & Sundberg, 1996; Tsai & Greer 2006; Yoon & Bennett, 2000).

According to several researchers, verbal development begins in the womb (Spence & DeCasper, 1987). Following birth, DeCasper and Fifer (1980) found that newborn infants prefer their mother's voices to those of other females and DeCasper and Spence (1987) found that newborn infants, two-three days old, demonstrated a preference for a passage that had been read to them by their mothers every day for six weeks prior to birth compared to a novel passage. Not only do infants prefer their mother's voices, they prefer familiar sounds. Theories about why this is so are related to conditioned reinforcement that begins prior to birth.

One explanation is that the sounds of the mother's voice are heard in utero and are paired with primary reinforcers present in the womb (e.g., warmth, nourishment, movement) and thus the mother's voice is conditioned as a reinforcer prior to birth. After birth, those pairings continue, with nourishment, touch, and the mother's face that is now paired with her voice. As a result of these pairings, the mother's face, and soon others, become conditioned reinforcers almost immediately after birth. Meltzoff and Moore (1983) found that newborn infants can imitate facial gestures within hours after birth, suggesting that it may also be the novelty of the face and the facial movements that act as primary reinforcers in selecting out the infants' observing and responding.

Conditioned reinforcement for listening to voices is also a necessary cusp for the development of both listener and speaker skills (Greer et al., 2011; Keohane, Luke, & Greer, 2008). Greer et al. (2011) conditioned voices as

reinforcers via a stimulus-stimulus pairing procedure. Results showed that all three participants' rate of learning accelerated, two children's observing responses increased, and two children's stereotypy decreased while their attention to a story read aloud by an adult increased. Keohane et al. (2008) implemented a rotated protocol package that included conditioning faces, voices, two- and three-dimensional stimuli, matching across the senses, and generalized imitation for three elementary students with ASD. Results of this treatment package demonstrated increased rates of learning and increased observing responses for all three children. While conditioning faces was one of the protocols implemented, it is impossible to isolate the effects of this intervention alone.

In the present study we used a conjugate stimulus-stimulus pairing procedure to condition adult faces and/or voices as reinforcers for four children with ASD. According to White (1971), conjugate reinforcement refers to "a schedule of reinforcement in which reinforcement is continuously present (e.g., the opportunity to eat) as long as a specified response is maintained at a criterion rate. Failure to maintain responding results in the discontinuance of reinforcement (e.g., the removal of the food dish) until responding again at criterion level" (p. 137). Conjugate reinforcement has resulted in conditioning novel stimuli as reinforcers (Cotter & Spradlin, 1971; Dunst, Storch, Hutto, & Snyder, 2007; Lindsley, 1956; Lovitt, 1968; Rovee & Rovee, 1969). We tested the effects of the conditioning procedure on the rate of acquisition of curricular objectives, emission of verbal operants, and observing responses to the presence of adults in the environment. In addition, we sought to determine whether adult faces or adult voices functioned as conditioned reinforcers, prior to or as a result of the intervention, so we conducted pre- and post-intervention probes of the reinforcing effects of adult faces and voices as well.

## Method

### *Participants*

We selected four males with developmental disabilities, ranging in age from four to eight years, based on classroom observations that indicated that adult faces and/or voices were not conditioned reinforcers. All participants emitted mands (requests) e.g., "I want jelly bean please" and tacts (object, event, or condition names) in complete sentences. All participants had a limited community of social reinforcers. In addition, the participants emitted low numbers of correct responses to learn units (response opportunities) across speaker and listener programs, low levels of observing responses, and low levels of verbal operants across three non-instructional settings, as confirmed by pre-intervention measures.

Participant A was a five-year-old male diagnosed with Autism Spectrum Disorder (ASD). Participant B was a four-year-old male diagnosed with a developmental disability (unspecified). Participant C was a five-year-old male diagnosed with Pervasive Developmental Disorder. Participant D was an 8-year-old male diagnosed with ASD.

### *Setting*

The experiment took place in a private publicly funded preschool 20 miles outside of a major city and a classroom in a public suburban elementary school 40 miles outside of a major city that implemented the CABAS® (Comprehensive Application of Behavior Analysis to Schooling, Greer, 2002) model. All instruction was conducted in the participants' classroom, as part of their normal classroom instruction. All pre- and post- intervention probe sessions and the intervention sessions were conducted outside of the participants' classrooms in an empty part of a hallway, where it was quiet and the walls were bare. The hallway setting contained a child-sized desk, a child-sized chair, and chairs for the experimenter and independent observer, when present.

For Participant D all pre- and post intervention sessions and the intervention were conducted in his home in a quiet room. The room consisted of a large dining table with eight chairs and bare walls. In addition, pre- and post observing response and verbal operant probe sessions were conducted in Participant D's classroom.

### *Materials*

During the conditioned reinforcement for listening to adult voices probe sessions the materials consisted of three child-sized chairs, a timer that counted forward, and two electronic Pal Pad (Adaptivation, Inc.) pressure-activated membrane switches connected to a tape recorder (see Figure 1). When a specific electronic switch was depressed, the tape recorder was activated and played a recorded voice reading a children's story. When the other electronic switch was depressed, no sound was emitted. During the stimulus-stimulus pairing procedure to condition faces as a reinforcer (the intervention) the materials consisted of two child-sized chairs, and a timer that counted forward. Pre and post-intervention measures of rate of acquisition of tacts required five sets of four 2-dimensional tact stimuli. These stimuli included pictures of animals (e.g., lizard, frog), flowers (e.g., lily, tulip), and musical instruments (e.g., trumpet, harp) printed in color, laminated, and affixed to 7.5cm X 12.5cm index cards. Other materials included data sheets and black pens to record the data.



*Figure 1.* The two electronic Pal Pad (Adaptivation, Inc.) pressure-activated membrane switches connected to a tape recorder used during the pre- and post-intervention probe sessions for voices as conditioned reinforcers. The switches were rotated periodically so that the participant had to find the switch that activated the voice recording.

### *Dependent Variables*

We tested the effects of conditioning adult faces as reinforcers on three dependent variables: 1) the rate of acquiring objectives across two broad curricular areas (listener responses and speaker responses), 2) observing responses to the presence of adults in the environment, and 3) verbal operants emitted across three non-instructional settings (i.e., lunch, art, and recess). To assess rate of learning, we took 1000 learn units consisting of listener responses (following 40 single-step instructions) and 1000 speaker learn units (20 tact stimuli) and divided them by the number of instructional objectives achieved. A learn unit consists of an instructional antecedent, the response from the child, and a consequence that functions to either reinforce future correct responses or a correction that functions to occasion future correct responses (Greer, 2002). The listener learn units included a total of 36 single-step commands (e.g., touch your toes, clap your hands) and nine “nonsense” commands (e.g., “la la la”). The 45 commands were grouped into nine sets of five commands each—four single-step instructions and one nonsense command. Each set was taught separately in blocks of 20 learn units, so that each command was presented four times per session.

In order to assess observing responses to the presence of adults in the environment, we measured whether the child oriented toward an adult (speaking or not speaking) across ten different scenarios (see Table 1). In order to assess the number of verbal operants emitted by the participants we conducted 10-min probes in three non-instructional settings (i.e., lunch, free play, and art) in which we measured the total number of mands, tacts, sequelics, and conversational units emitted (see Table 2 for a complete definition of each). In addition, as tests of the independent variable, we measured whether adult faces and voices functioned as conditioned reinforcers prior to and following the intervention.

**Table 1**  
*Verbal operants measured during pre- and post-intervention probes conducted during 10-min sessions across lunch, art, and recess.*

---

#### **Verbal Operants:**

**Mand:** A Mand specifies its reinforcer, and is produced in the presence of the item under deprivation without vocal antecedent and results in the delivery of the item

**Tact:** Production of a vocal response to a stimulus without vocal antecedent under the control of generalized social listener reinforcement

**Sequelic:** A verbal operant that occurs when an individual responds as a listener and speaker to intraverbals

**Conversational Unit:** An exchange that involves a listener and speaker in which each acts as speaker and listener to each other's intraverbals at least twice in an exchange

---

Table 2

*Observing Responses Measured during Pre- and Post-Intervention Probe Sessions**Antecedent and Opportunity for Observing Responses*

Participant orients toward a speaker when his name is called in a moderate, but detectable volume from a distance of 0.5-1.5 meters. "Orients" refers to the participant making eye contact or looking at the face of the experimenter or 3 <sup>rd</sup> party for a minimum of 1 s.
Participant orients toward a speaker when name is called in a moderate, but detectable volume from 1.5-2.5 meters.
Participant orients toward a speaker when the child is given a 1-step direction in a moderate, but detectable volume from 0.5-1.5 meters.
Participant orients toward a speaker when the child is given a 1-step direction in a moderate, but detectable volume from 1.5-2.5 meters.
Participant orients toward speaker when the child is spoken to in a moderate, but detectable volume from a distance of 0.5-1.5 meters.
Participant orients toward a speaker when another child is spoken to in a moderate, but detectable volume from 1.5-2.5 meters.
Participant orients toward an adult rearranging the child's materials on desk.
Participant orients toward an adult removing the child's materials from desk.
Participant orients toward an adult entering the room who is speaking in a moderate, but detectable volume.
Participant orients toward an adult entering room who is not speaking.

*Intervention: Face Conditioning*

We used a conjugate stimulus-stimulus pairing procedure during the intervention to condition adult faces as reinforcers. This procedure involved the experimenter getting the participant to orient to her face using non-vocal sounds (e.g., smacking lips, making loud kissing sounds) and not, for example, calling the child's name or saying, "look at me." Immediately upon the participant orienting toward and observing her face the experimenter delivered vocal, visual, and, in some cases, tactile reinforcement in the form of animated expressions, speaking, singing, and sometimes touching the face, head, or arms of the participant. If at any time the participant looked away from the experimenter's face for longer than one second then the experimenter ceased and attempted to regain the participant's attention to her face using non-vocal sounds. The definition of looking at the experimenter's face included the participant looking at any part of the experimenter's face (e.g., forehead, hair, eyes, cheeks, chin, mouth, etc.).

*Procedures and Data Collection*

*Learn Units-to-Criterion.* In the listener instruction, target commands and nonsense commands were selected and divided into sets of five responses each (four commands and one nonsense command). For each of the nine sets (five operants each) of listener learn units, the commands were presented four times each during a 20 learn unit session and each set was taught separately. The experimenter delivered the vocal antecedent, e.g., "clap your hands," without giving any visual cues. In other words, we ensured that the participant only responded to

the auditory properties of the antecedent. If the participant responded correctly to the command within three seconds, the experimenter delivered vocal praise or preferred edibles. For the single-step instructions, if the participant emitted an incorrect response or no response, the experimenter delivered a correction. Corrections involved re-presentations of the antecedent followed by an opportunity for the participant to respond again. In some cases, the experimenter provided a physical prompt in order for the participant to emit the correct responses. Correct responses that followed corrections were not reinforced. For the nonsense commands, the absence of any response was reinforced and incorrect responses were ignored: the experimenter paused and looked away for two seconds and then presented the next learn unit. Criterion consisted of the participants emitting 90% accuracy across two sessions or 100% accuracy for one session.

For tact (speaker responses) instruction each of the five sets of tacts was taught separately. The experimenter held up a stimulus and got the participant's attention. Once the participant looked at the stimulus he was given three seconds to emit the correct response, e.g., "harp." The experimenter delivered vocal praise and attention for correct responses and a correction for incorrect or no responses. For corrections, the experimenter re-presented the antecedent stimulus, provided the correct response, and gave the participant the opportunity to echo the correct response. Correct responses that followed corrections were not reinforced. Criterion consisted of the participants emitting 90% accuracy across two consecutive sessions or 100% accuracy in one session.

#### *Observing Response Probes*

During the pre and post-intervention observing response probe sessions, we measured the number of times out of ten opportunities that the participants looked at or in the direction of the approaching or speaking adult across the ten observing response scenarios (see Table 1), for a total of 100 response opportunities. We used different adults, both familiar and unfamiliar to the participants, and provided opportunities that were spaced to provide the most natural non-contrived setting. A plus (+) was recorded when the participant looked at or in the direction of the adult within one second and a minus (-) was recorded if the participant did not look at or in the direction of the adult within one second. When opportunities across all of the observing response scenarios were completed, the cumulative number of observing responses was tallied.

#### *Verbal Operant Probes*

During the pre and post-intervention verbal operant probe sessions, verbal operants were measured during three non-instructional settings (i.e., lunch, art, and recess). During these probe sessions two experimenters simultaneously but independently recorded the total number of verbal operants emitted by the participants during three 10-min sessions. At the end of each session, the verbal operants were tallied and categorized to derive a total number of each type (i.e., mands, tacts, sequelics, and conversational units).

*Tests of the Independent Variable—Faces and Voices as Conditioned Reinforcers*

*Probes for Adult Faces as Conditioned Reinforcers.* In order to assess conditioned reinforcement for observing adult faces, we conducted a 5-min probe using 5-s partial interval recording during which we measured whether the participant observed the experimenter's face while she moved her mouth and face in animated expressions without making vocal sounds. The experimenter mouthed the words to a poem or a passage from a book, for example, without using her voice. If at any moment in the 5-s interval the participant looked at or in the direction of the experimenter a plus (+) was recorded on the data sheet. If during the 5-s interval the participant did not look at or look in the direction of the experimenter a minus (-) was recorded on the data sheet. The participant was not required to observe the experimenter's face for the entire 5-s interval. At the end of the 5-min session, the number of pluses and minuses were tallied. In order for adult faces to be considered conditioned reinforcers the participant had to emit observing responses for a total of 45 intervals out of 60 (75%).

*Probes for Adult Voices as Conditioned Reinforcers.* During the conditioned reinforcement for listening to adult voices probe session the participant was required to depress a specified switch that activated an adult voice reading a children's story. If the participant depressed a second switch, no sound was emitted. The two switches were necessary in order to determine if the participant preferred listening to the voice on the recording; they were periodically rotated. Once the participant depressed the correct switch and activated the story, the experimenter started the countdown timer, which was set for five minutes. During the 5-s whole interval recording the experimenter recorded a plus (+) if the participant depressed the switch for the entire 5-s interval and recorded a minus (-) if the participant did not depress the switch for the entire 5-s interval. In order to control for passive depressing of the switch, the experimenter rotated the position of the switches after every ten intervals. The participant then needed to find the switch that again activated the voice recording. At the end of the 5-min session, the pluses and minus were tallied. In order for adult voices to be considered conditioned reinforcers for the participant, he had to depress the switch for a total of 45 intervals out of 60 (75%).

*Intervention.* During the intervention, we implemented the conjugate stimulus-stimulus pairing procedure. The first step was getting the participant to look at the experimenter's face. The experimenter used vocal sounds or musical instruments to get the participant's attention. For example, the experimenter chewed gum and blew bubbles that popped loudly, clicked her tongue, blew bubbles with her lips, stuck out her tongue and blew, rolled her tongue, or made sounds such as "da da da" or "la la la." At times she also played instruments such as a kazoo or harmonica. It is important to note that the experimenter emitted these sounds until the participant looked at her face. The sounds the experimenter produced were continuously changing, thus varying from moment to moment.

Once the participant oriented to the experimenter's face she immediately started a timer and she delivered two kinds of reinforcement, either separately or simultaneously, contingent upon the participant looking at her face. These included vocal reinforcement and/or tactile reinforcement.

Vocal reinforcement consisted of the experimenter singing animatedly or softly, reciting a nursery rhyme, or delivering vocal praise while the participant was looking at her face. The reinforcement was continually changing; the conjugate procedure involved novel sounds, facial expressions, and movements from the experimenter. As soon as the participant looked away for one second the experimenter stopped, and the trial ended. If the participant looked back at the experimenter within one second she continued with varied sounds and expressions. The experimenter made moment-to-moment decisions as to what sounds, expressions, or touches were reinforcing, and which ones the participant appeared not to like. Some participants preferred soft voices and touches, others preferred loud and exaggerated voices and expressions. Therefore, it was important for the experimenter to respond flexibly and determine immediately what was reinforcing for the participant in order to maintain the participant's observation of her face. The session continued until a total of 20 trials were completed or the participant met the criterion for the intervention. Once 20 trials were completed, the experimenter calculated the sum (e.g.,  $1+2+1+3+1+10+9+20+5+8+9+5+6+4+1+8+7+6+5+4=115$  cumulative s). The intervention continued until the participant emitted 160 cumulative s of observing the experiment's face across 20 trials. One session of the intervention was run each day.

### *Design*

We used a delayed non-concurrent probe design across participants in order to control for maturation and history. We conducted probes or gathered relevant learn unit data immediately prior to and following the intervention to condition adult faces and/or voices as reinforcers. Following the intervention, we repeated the probes and gathered the learn unit data for comparison with pre intervention data.

### *Interobserver agreement*

*Probes.* Interobserver agreement (IOA) was collected during all pre and post-intervention probe sessions for each participant. IOA for Participant A was conducted for 37% of sessions with a mean agreement of 94% with a range of 89-100%. IOA for B was conducted for 39% of sessions with a mean agreement of 95% and a range of 87-98%. IOA for Participant C was conducted for 34% of sessions with a mean agreement of 99% with a range of 95-100%. IOA for Participant D was conducted for 64% of sessions with a mean agreement of 97% with a range of 93-100%.

*Intervention.* During intervention, IOA was collected for Participants A, B, C, and D. IOA was calculated on a point-to-point basis for 100% of the intervention sessions for Participant A with a mean of 95% and a range of 87-98%. IOA was calculated on a point-to-point basis for 43% of the sessions for Participant B with a mean agreement of 99% and a range of 99-100%. IOA was calculated on a point-to-point basis for 57% of the intervention sessions for Participant C with a mean of 98% and a range of 94-100%. IOA was calculated on a point-to-point

basis for 50% of the intervention sessions for Participant D with a mean of 89% and a range of 85-93%.

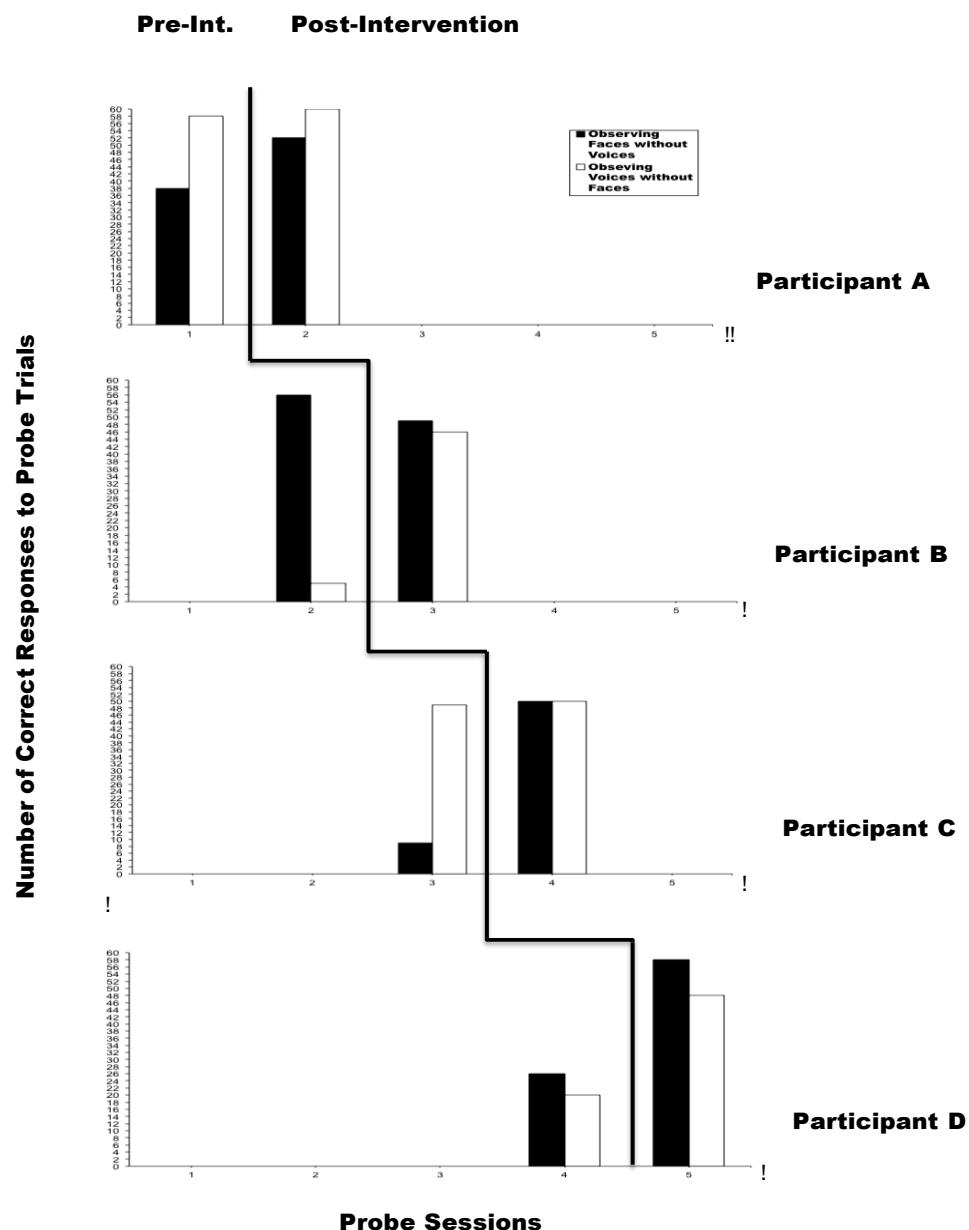
## Results

*Pre- and Post-Intervention Probes.* Figure 2 shows the results from the pre and post-intervention tests of conditioned reinforcement for observing human faces and conditioned reinforcement for listening to adult voices for Participants A, B, C, and D. Prior to intervention, Participant A demonstrated a total of 38 intervals out of 60 (5-min total probe with 5-s partial interval recording) for observing faces without voices. This is equivalent to saying that the participant looked at the experimenter during 63% of the intervals in 5-min probe session when the experimenter moved her face in animated ways or moved her lips while talking but without sound. During the conditioned reinforcement for human voices pre-intervention probe, Participant A emitted a total of 58 intervals out of 60 (5-min total probe with 5-s whole interval recording), or 97%. Therefore, prior to the conditioning intervention faces did not function as conditioned reinforcers for this participant, but voices did. Participant A required five sessions to meet criterion for observing faces during the intervention. Following the intervention, the conditioned reinforcement for observing human faces and conditioned reinforcement for listening to adult voices probes were repeated. The results showed an increase to 52 intervals out of 60 in total (87%) for observing faces without voices and 60 intervals out of 60 in total (100%) for listening to adult voices. Participant A demonstrated criterion-level responding to faces, thus indicating that the intervention functioned to condition faces as reinforcers.

Prior to the intervention, Participant B emitted a total of 56 intervals out of 60 (93%) during the conditioned reinforcement for observing faces probe and 5 intervals out of 60 (8%) during the listening to adult voices probe. Prior to the intervention, faces functioned as conditioned reinforcers for Participant B, but adult voices did not. Participant B required seven sessions to achieve the mastery criterion for observing faces and listening to voices during the intervention. Following the intervention, Participant B emitted 49 intervals out of 60 (82%) during the conditioned reinforcement for faces probe and 46 out of 60 intervals (77%) for probes of conditioned reinforcement for listening to voices. Participant B demonstrated criterion-level responding to voices probe, thus indicating that adult voices became conditioned reinforcers as a result of the intervention.

Prior to the intervention, Participant C emitted a total of 9 intervals out of 60 (15%) during the conditioned reinforcement for observing faces probe and 49 intervals out of 60 (82%) for the conditioned reinforcement for listening to adult voices probe. Therefore, adult faces did not function as conditioned reinforcers for Participant C, but listening to adult voices did. Participant C required seven sessions to achieve the mastery criterion for observing faces during the intervention. Following the intervention, Participant C emitted a total of 50 intervals out of 60 in total (83%) during the conditioned reinforcement for observing faces probe and he emitted a total of 50 intervals out of 60 (83%) during the conditioned reinforcement for listening to adult voices probe. The results indicated that the intervention functioned to condition faces as reinforcers for Participant C.

Prior to the intervention, Participant D emitted a total of 26 intervals out of 60 (43%) for the conditioned reinforcement for observing adult faces probe and 20 intervals out of 60 (33%) during the conditioned reinforcement for listening to voices probe. Participant D required four sessions to meet criterion for observing faces during the intervention.



*Figure 2.* Number of correct responses to pre-and post-intervention probe trials for test of conditioned reinforcement for faces and voices for Participants A-D. The solid black line indicates the intervention. Arrows indicate 0 responses.

Following the intervention, Participant D's observing responses increased to 58 intervals out of 60 in total (97%) for observing faces and 48 intervals out of 60 (80%) for listening to adult voices. Therefore, Participant D acquired both observing adult faces and listening to voices as conditioned reinforcers as a function of the intervention.

*Rate of Learning.* Figure 3 represents the rate of learning, as indicated by the number of learn units-to-criterion, for Participants A, B, C and D prior to and after the conditioning intervention. As a function of the conditioning intervention, Participant A's learn units-to-criterion for speaker operants decreased from 167 to 111. Participant A demonstrated a listener repertoire at the outset of the study. Participant B's learn units-to-criterion for speaker operants decreased from 143 to 111 as a function of the intervention. Participant B demonstrated a listener repertoire at the outset of the study. Participant C demonstrated a decrease from 333 to 100 learn units-to-criterion for listener responses and 143 to 200 learn units-to-criterion for speaker responses as a function of the conditioning intervention. Participant D's learn units-to-criterion decreased from 200 to 90 for speaker operants and 143 to 77 learn units-to-criterion for listener responses as a function of the intervention.

*Observing Responses.* Results from the observing responses probes indicated that Participant A's observing responses to the presence of adults increased from 24 to 39 out of a total of 100 opportunities following the intervention. Participants B's observing responses increased from 18 to 35 following the intervention, Participant C's observing responses increased from 12 to 48 following the intervention, and Participant D's observing responses increased from 32 during the pre-intervention probe to 100 out of a possible 100 opportunities following the intervention (Figure 4).

*Verbal Operant Probes.* Results from the verbal operant probes indicated that prior to the intervention, Participant A emitted a cumulative total of 4 mands, 13 tacts, 12 sequelics, and 0 conversational units across all three settings for a total duration of 30 minutes. During post-intervention probe, Participant A emitted 1 mand, 6 tacts, 9 sequelics, and 5 conversational units. Prior to the intervention, Participant B emitted 2 mands, 1 tact, and 0 sequelics and conversational units. During the post-intervention probe session, Participant B emitted 2 mands, 20 tacts, 4 sequelics, and 0 conversational units. During the pre-intervention probe Participant C emitted 4 mands, 0 tacts, 0 sequelics, and 0 conversational units. Following the intervention, he emitted 50 tacts, 46 mands, 19 sequelics, and 0 conversational units. Prior to the intervention Participant D emitted a cumulative total of 3 mands, 3 tacts, 1 sequelic, and 0 conversational units. During the post-intervention probe session, Participant D emitted a cumulative total of 12 mands, 1 tact, 0 sequelics, and 0 conversational units (Figure 5).

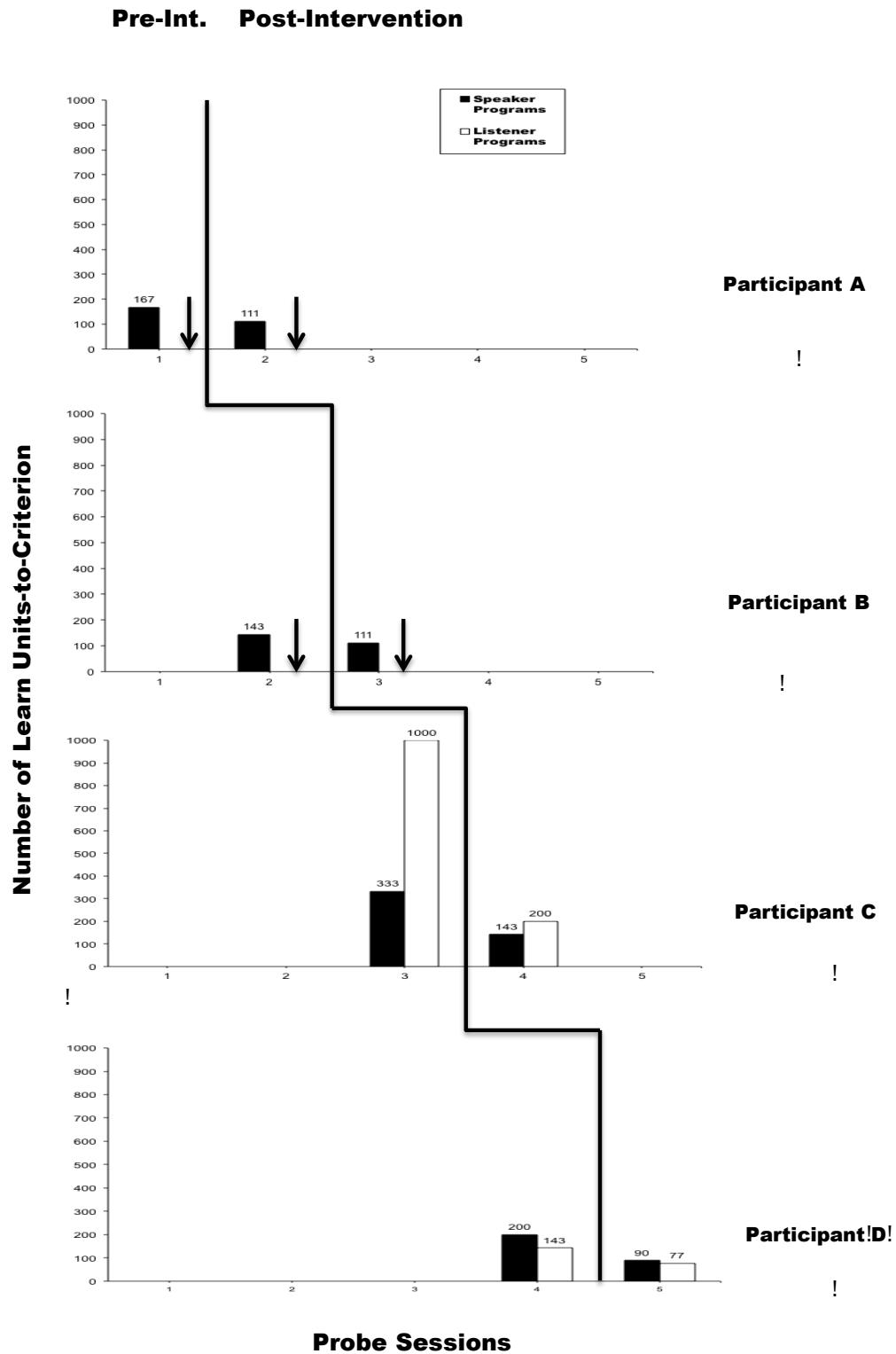


Figure 3. Number of pre-and post-intervention learn units-to-criterion for programs targeting speaker operants (tacts) for Participants A-D and listener operants for Participants C and D. The solid black line indicates the intervention. Arrows Indicate 0 responses

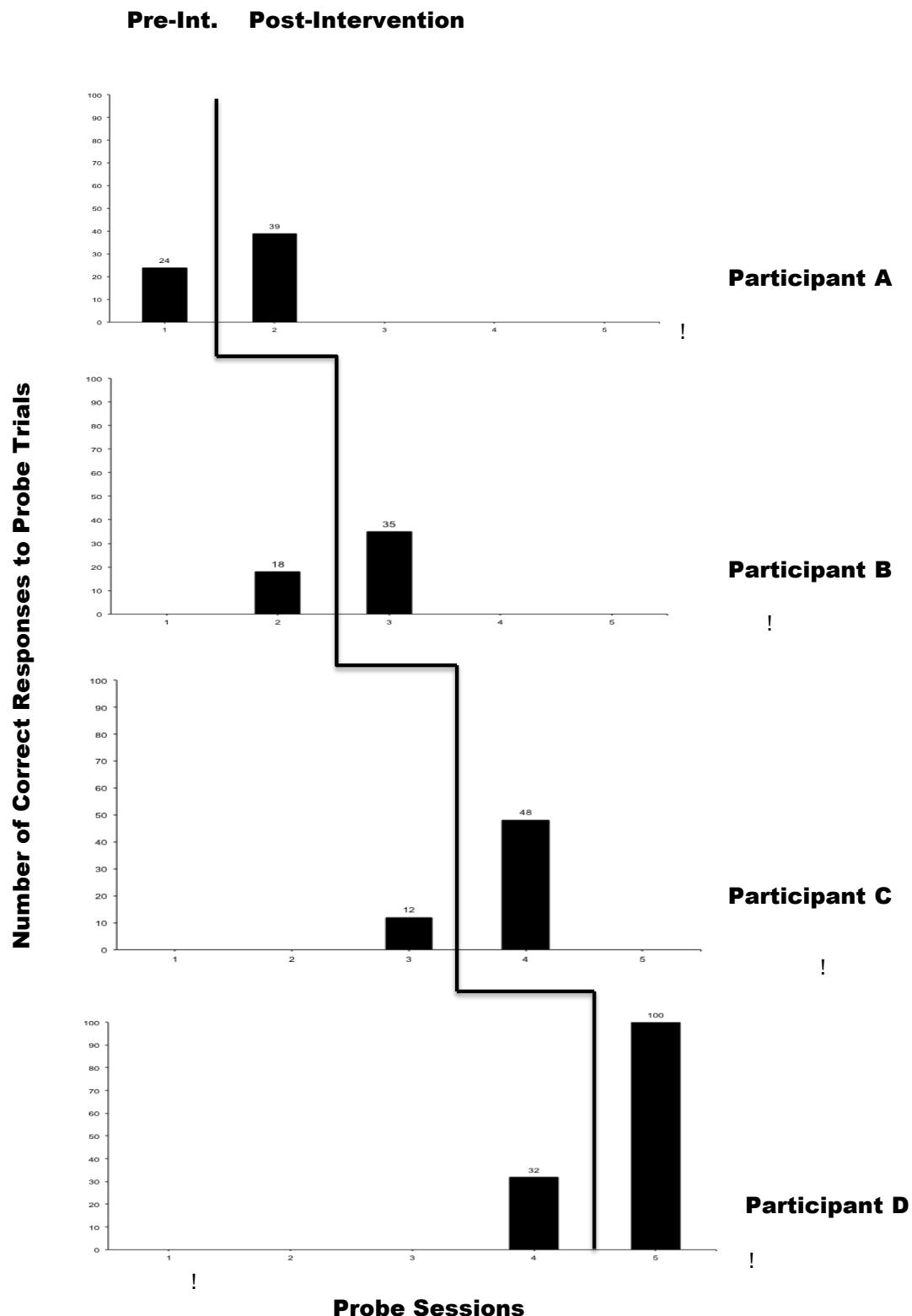
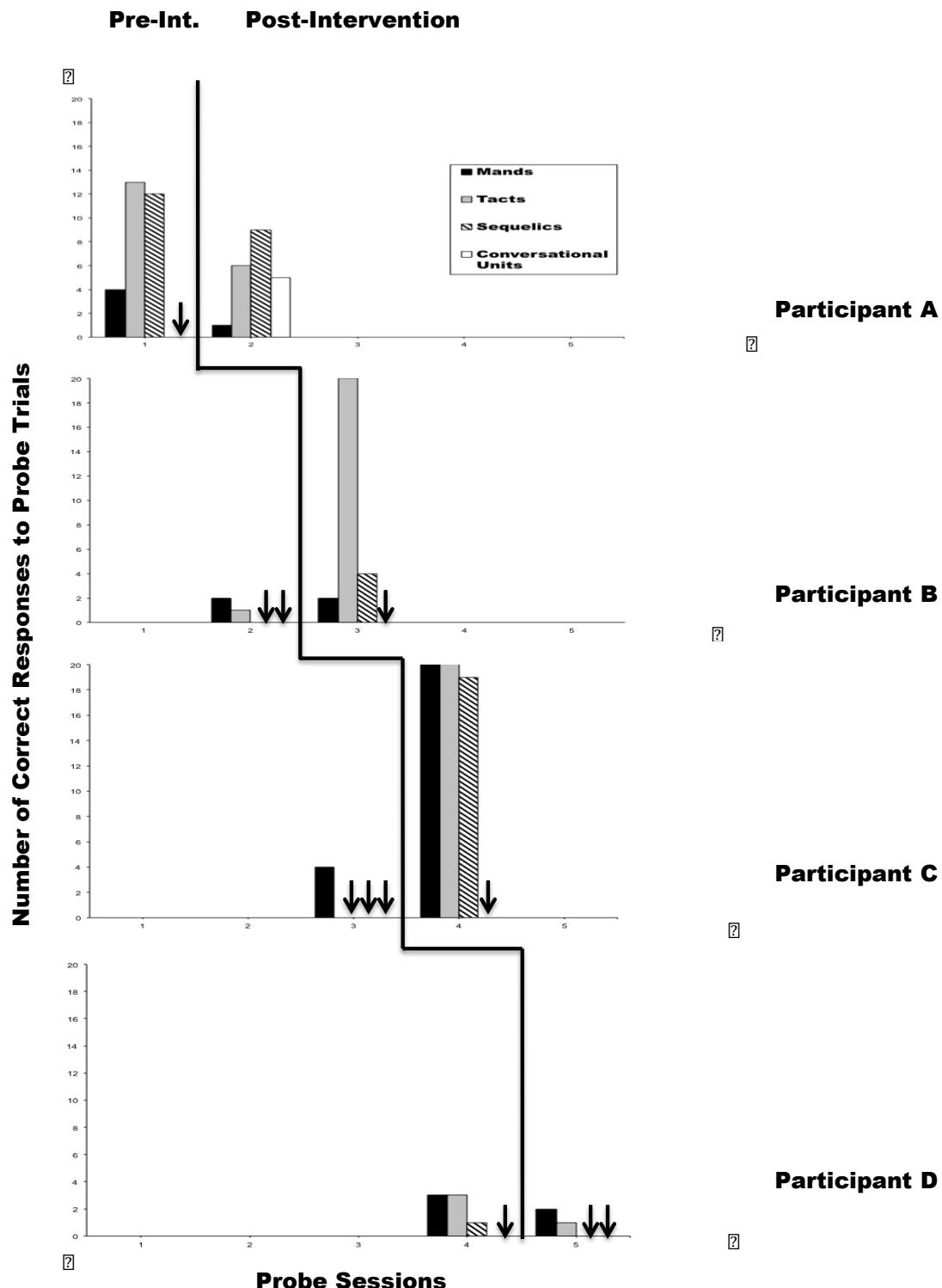


Figure 4. Number of correct responses to pre-and post-intervention probe trials for observing responses for Participants A-D—The Solid black line indicates the intervention. Arrows indicate 0 responses.



*Figure 5. Number of correct responses to pre-and post-intervention probe trials for mands, tacts, sequelics, and conversational units for Participants A-D. The solid black line indicates the intervention. Arrows indicate 0 responses.*

## Discussion

The results of the present study support the theory that the acquisition of conditioned reinforcement for observing the human face and/or listening to human voices are necessary pre-verbal developmental cusps, as theorized by the VBDT (Greer & Keohane, 2005; Greer & Ross, 2008; Greer & Speckman, 2009). As a result of the intervention, all four of our participants demonstrated 1) accelerated rates of learning, as measured by the number of learn units to criterion, 2) increases in the emission of tacts and mands (for Participants B, C, and D), as measured by the verbal operant probes, and, for three out of four of the participants, the emergence of higher order verbal operants (sequelics for Participants B and C and conversational units for Participant A), and 3) increased attention to the presence of the adults (speaking or not speaking), as measured by the observing response probes.

As a result of the acquisition of reinforcement for observing faces and/or voices, all four of our participants demonstrated increased attention to the presence of a potential speaker or listener, looked at a speaker more often, and listened and responded to instructional antecedents more readily, resulting in increased rates of learning across both listener (for Participants C and D) and speaker operants.

These findings are further supported by the results of the verbal operant probes, which were measures of social interaction. Tacts, sequelics, and conversational units are characteristically reinforced by a response from a listener. They are social repertoires, with social reinforcers. Mands, too, are mediated by a listener, but the reinforcer is the item or condition mandated. Participants A, B, and C all demonstrated increases in verbal operants that had social reinforcers. Although Participant A's post-intervention probes indicated that his number of mands, tacts, and sequelics decreased, conversational units emerged, which are higher order verbal operants. The individual alternates responding as both a speaker and a listener in a series of exchanges. This is a higher-order verbal operant. Participant D's tact and sequelic operants decreased following the intervention, but his mands increased. However, Participant D demonstrated 100% of observing responses in post-intervention probes, indicating that he was much more aware of the presence of adults in his environment.

The procedure we used during the intervention resulted in the conditioning of either one or two pre-verbal foundational cusps. Two of our participants (Participants A and C) had voices as conditioned reinforcers prior to the intervention, but faces did not function as reinforcers for observing. Both participants acquired faces as conditioned reinforcers as a function of the intervention. Participant B had faces as conditioned reinforcers prior to the intervention, but voices did not function as reinforcers. The intervention functioned to condition voices as reinforcers for Participant B. For Participant D, neither faces nor voices functioned as reinforcers prior to the intervention. Post-intervention results indicated that both of these cusps were established for Participant D.

We attribute these results, particularly the conditioning of voices in addition to faces as reinforcers, to the ever-changing, novel, conjugate stimulus-stimulus pairings that occurred during the intervention. The moment-to-moment

responsiveness by the experimenter to the participant was critical. In some cases the experimenter's loud tone appeared aversive to the participant, so she lowered her voice and spoke in softer tones while in other cases the participant seemed to prefer louder, more exaggerated tones. The experimenter's continual changes in response to the participant's observed reactions to her voice and facial expressions, animations, and/or tactile touch likely led to the success of this procedure. Consequently, those who will implement this procedure in the future need to constantly observe and respond to the participant and adjust their volume, proximity, intensity, and/or touch in order to ensure that the procedure is in fact conditioning the face and/or voice as a reinforcer.

In summary these results suggest that the acquisition of the human face and/or human voice as conditioned reinforcers are critical in language development and social behavior. Establishing these two pre-verbal cusps is the foundation to the development of language, and their importance is indicated for children with ASD.

## References

Adaptivation, Inc. Sioux Falls, SD: [http://adaptivation.com/product\\_detail.php?ID=44](http://adaptivation.com/product_detail.php?ID=44)

Arnold, A., Semple, R. J., Beale, I., & Fletcher-Flinn, C. M. (2000). Eye contact in children's social interactions: What is normal behaviour?. *Journal of Intellectual and Developmental Disability*, 25, 207-216. doi:10.1080/13269780050144271

Baron-Cohen, S., Baldwin, D. A., & Crowson, M. (1997). Do children with autism use the speaker's direction of gaze strategy to crack the code of language? *Child Development*, 68, 48-57. doi: 10.2307/1131924

Cleveland, A., Kobiella, A. & Striano, T. (2006). Intention or expression? Four month-olds' reactions to a sudden still-face. *Infant Behavior and Development*, 29, 299-307. doi: 10.1016/j.infbeh.2005.12.004

Cotter, V. W., & Spradlin, J. E. (1971). A nonverbal technique for studying music preference. *Journal of Experimental Child Psychology*, 11, 357-365. doi: 10.1016/0022-0965(71)90040-3

DeCasper, A. J. and Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, 208, (4448), 1174-1176. doi: 10.1126/science.7375928

DeCasper, A. J., & Spence, M. J. (1987). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development*, 9, 133-150. doi: 10.1016/0163-6383(86)90025-1

Dinsmoor, J. (1983). Observing and conditioned reinforcement. *Behavior and Brain Sciences*, 6, 693-704. doi: 10.1017/S0140525X00017969

Dunst, C. J., Storch, A. J., Hutto, M. D., & Snyder, D. (2007). Relative effectiveness of episodic and conjugate reinforcement on child operant learning. *Winterberry Research Syntheses*, 1, 1-15.

Ellsworth, C. P., Muir, D., & Hains, S. M. J. (1993). Social competence and person-object differentiation: An analysis of the still-face effect. *Developmental Psychology*, 29, 63-73. doi: 10.1037/0012-1649.29.1.63

Greer, R. D. (2002). *Designing teaching strategies: An applied behavior analysis*

system approach. New York, U.S.A.: Academic Press.

Greer, R. D., Becker, B. J., Saxe, C. D., & Mirabella, R. F. (1985). Conditioning histories and setting stimuli controlling engagement in stereotypy or toy play. *Analysis and Intervention in Developmental Disabilities*, 5, 269-284. doi: 10.1016/0270-4684(85)90015-1

Greer, R. D., Dorow, L. G., Wachhaus, G., & White, E. R. (1973). Adult approval and students' music selection behavior. *Journal of Research in Music Education*, 21, 345-354. doi: 10.2307/3344908

Greer, R. D., & Keohane, D. D. (2005). The evolutions of verbal behavior in children. *Behavioral Development Bulletin*, 1, 31-47. Reprinted in 2006 in the *Journal of Speech and Language Pathology: Applied Behavior Analysis*, 1, 111-140. Retrieved from: <http://www.baojournal.com/SLP-ABA%20WEBSITE/SLP-ABA-VOL-1/SLP-ABA-1-2.pdf>

Greer, R. D., & Ross, D. E. (2008). *Verbal behavior analysis: Inducing new capabilities and expanding communication repertoires for children*. Boston, MA: Allyn & Bacon.

Greer, R. D., & Speckman, J. (2009). The integration of speaker and listener responses: A theory of verbal development. *The Psychological Record*, 59, 449-488.

Greer, R. D., Pistoljevic, N., Cahill, C., & Du, L. (2011). Conditioning voices as reinforcers for listening responses in preschoolers with autism. *The Analysis of Verbal Behavior*, 27, 103-124.

Hains, S. M., & Muir, D. W. (1996). Infant sensitivity to adult eye direction. *Child Development*, 67, 1940-1951. doi: 10.2307/1131602

Haith, M. M., Bergman, T., & Moore, M. J. (1977). Eye contact and face scanning in early infancy. *Science*, 198, 853-855. doi: 10.1126/science.918670

Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature*, 504, 427-431. doi: 10.1038/nature12715

Keohane, D. D., Luke, N. & Greer, R. D. (2008). The things we care to see: The effects of rotated protocol immersion on the emergence of early observing responses. *Journal of Early Intensive Behavior Intervention*, 5(1), 23-39. Retrieved from: <http://www.baojournal.com/JEIBI/JEIBI-VOL-5/JEIBI-5-1.pdf>

Keohane, D. D., Pereira Delgado, J. A., & Greer, R. D. (2009). Observing responses: Foundations of higher order verbal operants. In R. A. Rehfeldt, & Barnes-Holmes, (Eds.), *Derived relational responding: Applications for learners with autism and other developmental disabilities* (pp. 41-62). Oakland, CA: New Harbinger Publications.

Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin*, 100, 78-100. doi: 10.1037/0033-2909.100.1.78

Lindsley, O. R. (1956). Operant conditioning methods applied to research in Chronic schizophrenia. *Psychiatric Research Reports*, 5, 118-139.

Longano, J., & Greer, R. D. (2006). The effects of a stimulus-stimulus pairing procedure on the acquisition of conditioned reinforcement for observing and manipulating stimuli by young children with autism. *Journal of Early and Intensive Behavior Interventions*, 3, 135-150. Retrieved from:

http://www.baojournal.com/JEIBI/JEIBI-VOL-3/JEIBI-3-1.pdf

Lovitt (1968). Free-operant assessment of musical preference. *Journal of Experimental Child Psychology*, 6, 361-367. doi: 10.1016/0022-0965(68)90117-3

Massaro, D. W., & Bosseler, A. (2006). Read my lips. *Autism*, 10, 495-510. doi:10.1177/1362361306066599

Meltzoff, A. N., & Moore, M. K. (1983). Newborn infants imitate adult facial gestures. *Child Development*, 54, 702-709. doi: 10.2307/1130058

Miguel, C. F., Carr, J. E., & Michael, J. (2002). The effects of a stimulus-stimulus pairing procedure on the vocal behavior of children diagnosed with autism. *The Analysis of Verbal Behavior*, 18, 3-13.

Mirenda, P. L., Donnellan, A. M., & Yoder, D. E. (1983). Gaze behavior: A new look at an old problem. *Journal of Autism and Developmental Disorders*, 13, 397-409.doi:10.1007/BF01531588

Nuzzolo-Gomez, R., Leonard, M. A., Ortiz, E., Rivera, C. M., & Greer, R. D. (2002). Teaching children with autism to prefer books or toys over stereotypy or passivity. *Journal of Positive Behavior Interventions*, 4, 80-87. doi: 10.1177/109830070200400203

Pereira-Delgado, J., Greer, R. D., Speckman, J., & Goswami, A. (2009). Effects of conditioning reinforcement for print stimuli on match-to-sample responding in preschoolers. *Journal of Speech Language Pathology and Applied Behavior Analysis*, 3.2-3.3, 60-77. Retrieved from http://www.baojournal.com/SLP-ABA%20WEBSITE/SLP-ABA-VOL-3/SLP-ABA-Consolidated-3-2-3-3.pdf

Rheingold, H. L., Gewirtz, J. L., & Ross, H. W. (1959). Social conditioning of vocalizations in the infant. *Journal of Comparative and General Psychology*, 52, 68-73.

Rosales-Ruiz, J., & Baer, D. M. (1997). Behavioral cusps: A developmental and pragmatic concept for behavior analysis. *Journal of Applied Behavior Analysis*, 30(3), 533-544. doi:10.1901/jaba.1997.30-533

Rovee, C. K., & Rovee, D. T. (1969). Conjugate reinforcement of infant exploratorybehavior. *Journal of Experimental Child Psychology*, 8, 33-39. doi:10.1016/0022-0965(69)90025-3

Senju, A., Kikuchi, Y., Hasegawa, T., Tojo, Y., & Osanai, H. (2008). Is anyone looking at me? Direct gaze detection in children with and without autism. *Brain and Cognition*, 67, 127-139. doi:10.1016/j.bandc.2007.12.001

Skinner, B. F. (1957). *Verbal Behavior*. Acton, MA: Copley Publishing Group and the B. F. Skinner Foundation.

Smith, R., Michael, J., & Sundberg, M. L. (1996). Automatic reinforcement and automatic punishment in infant vocal behavior. *Analysis of Verbal Behavior*, 13, 39-48.

Spence, M., & DeCasper, A. J. (1987). Prenatal experience with low-frequency maternal-voice sounds influence neonatal perception of maternal voice samples. *Infant Behavior and Development*, 10, 133-142. doi:10.1016/0163-6383(87)90028-2

Striano, T., & Bertin, E. (2004). Contribution of facial and vocal cues in the still-

Face-response of 4-month-old infants. *Infant Behavior & Development*, 27, 499-508. doi:10.1016/j.infbeh.2004.06.002

Sundberg, M. L., Michael, J., Partington, J. W., & Sundberg, C. A. (1996). The role of automatic reinforcement in early language acquisition. *The Analysis of Verbal Behavior*, 13, 21-37.

Symons, L. A., Hains, S. M. J., & Muir, D. W. (1998). Look at me: Five -month-old infants' sensitivity to very small deviations in eye-gaze during social interactions. *Infant Behavior & Development*, 21, 531-537. doi: 10.1016/S0163-6383(98)90026-1

Tsai, H., & Greer, R. D. (2006). Conditioned preference for books and faster acquisition of textual responses by preschool children. *Journal of Early and Intensive Behavioral Interventions* 3, 35-60. Retrieved from <http://www.baojournal.com/JEIBI/JEIBI-VOL-3/JEIBI-3-1.pdf>

White, O. (1971). *A glossary of behavioral terminology*. Champaign, Illinois: Research Press Company.

Yoon, S. Y., & Bennett, G. (2000). Effects of a stimulus-stimulus pairing procedure on conditioning vocal sounds as reinforces. *Analysis of Verbal Behavior*, 17, 75- 88.