

*TEACHING EMPATHY SKILLS TO CHILDREN WITH AUTISM*

JESSICA A. SCHRANDT

QUEENS COLLEGE AND THE GRADUATE CENTER OF THE CITY  
UNIVERSITY OF NEW YORK

DAWN BUFFINGTON TOWNSEND

INSTITUTE FOR EDUCATIONAL ACHIEVEMENT

AND

CLAIRE L. POULSON

QUEENS COLLEGE AND THE GRADUATE CENTER OF THE CITY  
UNIVERSITY OF NEW YORK

The purpose of this study was to teach empathetic responding to 4 children with autism. Instructors presented vignettes with dolls and puppets demonstrating various types of affect and used prompt delay, modeling, manual prompts, behavioral rehearsals, and reinforcement to teach participants to perform empathy responses. Increases in empathetic responding occurred systematically with the introduction of treatment across all participants and response categories. Furthermore, responding generalized from training to nontraining probe stimuli for all participants. Generalization occurred from dolls and puppets to actual people in a nontraining setting for 2 participants. Generalization was observed initially to the nontraining people and setting for the other participants, but responding subsequently decreased to baseline levels. Introduction of treatment in this setting produced rapid acquisition of target skills.

DESCRIPTORS: autism, empathy, social skills

Deficits in empathy and perspective taking have been well documented in children with autism (e.g., Rutter, 1978; Sigman, Kasari, Kwon, & Yirmiya, 1992; Yirmiya, Sigman, Kasari, & Mundy, 1992). For example, Yirmiya et al. demonstrated that after watching video segments of children experiencing different events and emotions, high-functioning children with autism were less able than their typically developing peers to label others' emotional

states, take the perspective of another person, and respond with empathy.

Empathy and perspective taking serve an important role in what Rheingold and Hay (1980) called *prosocial behavior* (e.g., helping, sharing, turn taking). Eisenberg (1992) and Rheingold, Hay, and West (1976) observed prosocial behavior, such as sharing and giving comfort, in typically developing children as young as 18 months old. Typically developing children display perspective-taking skills around 4 years of age (MacNamara, Baker, & Olson, 1976; Wimmer & Perner, 1983). It is easy to see why such behavior is important in the development of social relationships. Peers and family members may initiate social interaction more frequently with children who take the perspective of and demonstrate empathy towards others.

Although many researchers have documented deficits in empathy and perspective-taking skills

---

We thank Alana Bellizzi, Casey Gaetano, Julia Smith, and Jennifer Williams (teachers at the Institute for Educational Achievement), and Paul Argott, all of whom assisted in carrying out this study. Special thanks to the participants and their families for making this study possible.

Correspondence concerning this article should be addressed to Jessica Schrandt, 4048 Harriet Ave., Minneapolis, Minnesota 55409 (e-mail: jessyschrandt@yahoo.com).

doi: 10.1901/jaba.2009.42-17

in individuals with autism (e.g., Baron-Cohen, Leslie, & Frith, 1985; Bauminger & Kasari, 1999; Sigman *et al.*, 1992; Yirmiya *et al.*, 1992), fewer have focused on the extent to which these and related social skills can be taught. In the studies reviewed, modeling, prompting, and reinforcement were the most frequently used and effective procedures for teaching social skills.

Gena, Krantz, McClannahan, and Poulson (1996) used modeling, prompting, and reinforcement to teach 4 children with autism contextually appropriate affective responding to various scenarios (e.g., showing sympathy, laughing about absurdities). The treatment resulted in increased appropriate affect across participants and response categories, and these skills generalized across nontraining scenarios, therapists, time, and settings.

Similarly, Harris, Handleman, and Alessandri (1990) used prompting and reinforcement to teach 3 adolescent boys to display helping responses in a variety of situations. Although responding generalized from training to nontraining people and settings to some extent for each of the boys, levels of generalized responding were lower than for training responding.

Expanding on the Harris *et al.* (1990) study, Reeve, Reeve, Townsend, and Poulson (2007) used video modeling, prompting, and reinforcement to teach 4 children with autism to emit helping responses in the presence of multiple discriminative stimuli from a variety of helping categories (e.g., replacing broken materials, locating objects, putting away items). The treatment package was effective in teaching helping responses during training trials, and responding generalized across nontraining stimuli and response categories.

The aforementioned studies, like the present study, approached deficits in social behavior as a stimulus control problem. Many behavioral deficits in individuals with autism have been attributed to the phenomenon of stimulus overselectivity, the tendency to respond to a

limited number of cues when presented with complex stimuli (Bailey, 1981; Lovaas, Koegel, & Schreibman, 1979; Schreibman, Koegel, & Craig, 1977). When learning new discriminations, individuals with autism may attend to few or irrelevant stimuli in the environment. Thus, failures in empathy (or other complex social behavior) may not reflect a deficit of necessary responses in the repertoire but rather an inability to differentiate the stimuli in the presence of which specific responses (e.g., offering assistance, demonstrating interest) would be appropriate.

In the natural environment, discriminative stimuli for empathetic responding are both varied and complex. In previous studies, researchers effectively promoted generalization of social skills from training to nontraining stimuli by (a) training with multiple exemplars within a stimulus class (e.g., showing a variety of reasons for sadness), and (b) training with discriminative stimulus compounds (e.g., presenting a combination of vocal and nonvocal "sadness" stimuli).

In the present study, we sought to examine whether procedures used previously to teach other social skills could be extended to teaching empathy skills to children with autism. Specifically, the purpose of the present study was to assess the extent to which a treatment package consisting of the presentation of affective discriminative stimulus compounds, prompt delay, modeling through auditory scripts, manual prompts, behavioral rehearsals, and reinforcement was effective in teaching empathy skills in a pretend-play setting to children with autism. Behavioral rehearsals and prompt delay (Charlop, Schreibman, & Thibodeau, 1985; Halle, Marshall, & Spradlin, 1979) were used to prevent and correct errors, which were judged to be a potential difficulty. We also assessed the extent to which empathy skills generalized from training to nontraining probe stimuli and from training dolls and puppets to actual people in a nontraining setting.

## METHOD

*Participants*

Participants were 4 children with autism who attended an education program offering center-based and in-home behavioral intervention for children with autism for 5.5 hr per day, 5 days per week. They were identified for participation based on anecdotal reports from their teachers that they did not regularly demonstrate empathy toward others. All participants engaged in minimal stereotypic or disruptive behavior and had experience with token motivational systems. In addition, all participants had prerequisite skills in vocal imitation of three-word phrases modeled by an instructor and on auditory recordings.

Josh was 4 years 5 months old at the start of the study. He obtained a standard score of 114 (age equivalent: 5 years 7 months) on the Peabody Picture Vocabulary Test (3rd ed., PPVT-III) and 106 (age equivalent: 4 years 10 months) on the Expressive Vocabulary Test (EVT). In addition to the behavioral intervention program, he also attended an integrated preschool classroom for several hours per day, spoke in full sentences with clear articulation, and established eye contact regularly during social interaction. Although he sometimes made comments when observing displays of affect by others (e.g., “crying,” “sad”), he rarely displayed appropriate empathy responses.

Jacob was 6 years old at the start of the study. He obtained a standard score of 61 (age equivalent: 2 years 9 months) on the PPVT-III and 41 (age equivalent: 2 years 4 months) on the EVT. His language consisted mostly of requests, he rarely initiated social interaction, and his eye contact and attending to others were often poor. He demonstrated no empathy responses to displays of affect by others.

Luke was 5 years 6 months old at the start of the study. He obtained a standard score of 56 (age equivalent: 1 year 10 months) on the PPVT-III and 55 (age equivalent: 2 years 7 months) on the EVT. His language consisted

mostly of single words and short phrases, and his social initiations were usually nonvocal (e.g., smiling, walking toward a person). Like Jacob, he did not demonstrate empathy responses to displays of affect by others.

Ali was 8 years 9 months old at the start of the study. She obtained a standard score of 40 (age equivalent: 3 years 0 months) on the PPVT-III and 41 (age equivalent: 3 years 11 months) on the EVT. Her language consisted mostly of single words and short phrases. She rarely demonstrated empathy skills, although she sometimes said “sorry” when she accidentally bumped into another person.

*Setting and Materials*

Sessions were conducted in a small room furnished with small tables, chairs, and bookshelves. Toys were stored on the shelves, and dolls and puppets used for training and probe trials were placed in separate bins in front of the shelves. A Language Master was placed on one of the tables. A digital video camera was placed in one corner of the room and was used to record some sessions.

*Vignettes and Target Responses*

Affective discriminative stimulus compounds contained one motor and one vocal component. They were presented in brief vignettes in which an instructor held up a doll or puppet and pretended to make it say a statement paired with an action. Vignettes fell into three categories: sadness or pain, happiness or excitement, and frustration. For example, in the sadness or pain category, a vignette consisted of the instructor holding a puppet, making it bump the table, and saying “ouch!” as though the puppet had hurt its leg. Vignettes were assigned randomly to training and probe conditions. Sets of training and probe dolls and puppets were matched based on similarity (e.g., each set contained boys, girls, and animals). To promote generalization from training to non-training stimuli, dolls and puppets were not assigned to specific vignettes.

Table 1

Training and Probe Discriminative Stimuli and Training Responses for the Sadness or Pain Category

Discriminative stimuli	Responses
Training	
“Ouch!” and bumping leg on table	“Are you okay?” and patting arm
“I don’t feel good” and sitting down	“Are you alright?” and patting arm
“I hurt my elbow” and touching elbow	“It’s okay” and patting arm
“My stomach hurts” and rubbing stomach	
“Oww!” and falling down	
“I’m so upset” and sniffing or crying	
“I burned my hand!” and grabbing toast and dropping it	
Generalization	
“Oh no!” and wiping eyes with tissue	
“I cut my finger!” and cutting with scissors	
“I have a headache” and holding forehead	

*Empathy* was operationally defined as a contextually appropriate response to a display of affect by a doll, puppet, or person that contained motor and vocal components (in any order) and began within 3 s of the end of the display. For the previous example, the participant could be taught to say, “Are you okay?” and to pat the puppet’s arm. One motor response and three different vocal responses were taught for each response category. To promote generality to nontraining vignettes, the responses were designed such that any response taught for a particular vignette was appropriate for any vignette in the same response category. In addition, other contextually appropriate responses not targeted during training were scored as correct. For example, in the sadness or pain category, saying, “That’s too bad” while giving the doll a hug was scored correct even though this response was not targeted directly. The same response was not, however, scored correct for the happiness or excitement category, because it was not contextually appropriate. A complete list of vignettes and target responses for each response category is provided in Tables 1, 2, and 3.

### *Procedure*

*General procedure.* Sessions were conducted four to five times per week and lasted approximately 20 to 30 min. A trial began when the instructor presented a vignette with a

doll or puppet. The prompter, seated behind the participant, waited for a designated interval for a response and then delivered a consequence (following training trials) or gave no feedback (following baseline and probe trials). A baseline or probe trial ended when 3 s elapsed. Although 3 of the participants received training in only the sadness or pain response category, baseline trials for the other two categories were presented during each session. Dolls, puppets, toys, and a token motivational system were placed on the table between the participant and the instructor. Prior to each session, participants chose from an array of preferred snacks and activities that were delivered in exchange for every 10 tokens earned throughout the session (approximately three times per session).

*Baseline.* Throughout the session, the prompter delivered one token per trial for appropriate sitting and attending to the vignettes. Tokens were delivered randomly either between trials as the instructor set up materials for the next vignette or directly before the instructor presented the vignette. Instructors and prompters did not deliver explicit instructions or prompts to the participant regarding the dolls, puppets, or toys on the table. Correct responses were followed by a conversational exchange by the instructor (e.g., “thank you”). No feedback was given for incorrect responses.

*Treatment package for training sessions.* Training sessions consisted of 30 trials (seven training

Table 2

Training and Probe Discriminative Stimuli and Training Responses for the Happiness or Excitement Category

Discriminative stimuli	Responses
Training	
"This is fun!" and pushing car	"Can I see?" and moving hand toward own body
"I finished!" and coloring picture	"Show me!" and moving hand toward own body
"I did it!" and completing puzzle	"Let me see!" and moving hand toward own body
"Watch me!" and stretching Slinky	
"Look at this!" and showing gift box	
"I won!" and playing Hungry Hungry Hippos and putting both hands in the air	
"I found something!" and showing lizard	
Generalization	
"I love this!" and rolling Animaniacs ball	
"Look what I got!" and showing Magic Wheel	
"This is the best!" and playing frog piano	

and three nontraining probe-stimulus trials per response category) presented in random order. The prompter delivered manual and auditory prompts according to a prompt-delay sequence (Charlop et al., 1985; Halle et al., 1979) to prevent errors. Specifically, the prompter used manual and auditory prompts immediately (0-s delay) for three sessions when training was introduced for a particular response category. During the fourth session, the prompter used a 1-s delay. The prompter increased the delay length by 1 s with a maximum of a 3-s delay following each session in which the participant responded correctly for at least four of the seven trials in a particular response category. The

prompter decreased the delay by 1 s if the participant responded correctly on fewer than four of the seven training trials. During training trials, the instructor delivered tokens and behavior-specific praise following all correct unprompted responses.

A behavioral rehearsal sequence was used for all training trials in the 0-s delay condition and when the participant did not respond or responded incorrectly in all other prompt-delay conditions. First, the prompter manually prompted a correct motor response and simultaneously played an auditory script on a Language Master, a device that plays aloud phrases recorded on a strip of magnetic tape

Table 3

Training and Probe Discriminative Stimuli and Training Responses for the Frustration Category

Discriminative stimuli	Responses
Training	
"This won't work!" and putting Sylvester car on table	"I can help you" and reaching one hand toward puppet/person, palm turned slightly up
"I can't do it!" and trying to put shape in sorter	"Want some help?" and reaching one hand toward puppet or person, palm turned slightly up
"I can't get it" and trying to zip up coat	"I'll help you" and reaching one hand toward puppet or person, palm turned slightly up
"Oh man!" and dropping bin of blocks	
"It's too hard!" and trying to put K'Nex pieces together	
"It's stuck!" and trying to remove lid from box	
"I broke it!" and holding up broken Lego model	
Generalization	
"It's not working!" and trying to color with marker with cap on	
"My arm is stuck!" and trying to put on sweater	
"This is too tight!" and trying to open jar	

along the bottom of a card that is pushed through it. If the participant did not imitate the auditory script, the prompter played the card again and waited for a response. If the participant did not imitate the auditory script after two presentations, the prompter provided vocal prompts (e.g., "Say 'It's okay'") until the participant correctly imitated the script. Then, the prompter played the auditory script again and repeated this sequence until the participant correctly imitated the auditory script immediately after it was played on the Language Master. Once the participant imitated the auditory script without prompts, the instructor presented the same vignette, and the prompter partially prompted the vocal and motor responses (e.g., played the first word of the auditory script and lightly tapped the participant's arm). If the participant responded correctly, the prompter then gave the participant an opportunity to respond to the same vignette without prompts. If the participant did not respond or an error occurred, the sequence was repeated until the participant emitted a correct independent response.

*Generalization to nontraining probe stimuli.* These trials were randomly interspersed throughout all training sessions (including baseline sessions); consisted of vignettes similar to those used during training trials; and incorporated dolls, puppets, and toys never associated with training. During these trials, instructors and prompters followed the same procedures used during the baseline condition.

*Generalization to nontraining people and setting.* These sessions occurred approximately once per week during baseline and training. Sessions consisted of seven trials per response category. An actual person replaced the dolls and puppets and presented the training vignettes used in the training setting. Sessions were conducted in a large conference room not associated with training. Again, instructors and prompters followed the same procedures used during the baseline condition.

### *Experimental Design*

A multiple baseline design across participants was used to assess the extent to which the treatment package was effective in teaching empathetic responding to displays of sadness or pain to 4 participants (Josh, Jacob, Luke, and Ali). For 1 participant (Josh), a second multiple baseline design across response categories was used to assess the extent to which the treatment package was effective in teaching empathetic responding across three response categories (sadness or pain, happiness or excitement, and frustration).

### *Data Collection and Interobserver Agreement*

The experimenter and one of four instructors conducted interobserver agreement assessments for at least 35% of sessions in each of the experimental conditions for each participant. Observers met a criterion of 80% agreement in data collection before agreement measures began. Independent observers sat at separate tables and recorded motor and vocal responses verbatim. For each trial, they circled "yes" for responses that contained both a motor and vocal component, were contextually appropriate, and began within 3 s of the end of the vignette, and "no" for responses that did not meet these requirements. Agreements were defined as trials in which both observers circled "yes" or "no," and disagreements were defined as trials in which one observer circled "yes" and the other circled "no." Agreements were divided by agreements plus disagreements, and this ratio was converted to a percentage. Across all participants and experimental conditions, interobserver agreement ranged from 83% to 100%, indicating no more than one disagreement per response category in a single session.

### *Procedural Integrity*

To ensure the integrity of the independent variable, procedural integrity measures were collected during all interobserver agreement sessions. Independent observers assessed accurate presentation of auditory and manual

prompts as well as accurate delivery of tokens for each trial during all conditions. Prompt delivery was considered accurate if prompts were delivered only after the designated amount of time during the prompt-delay condition or when participants responded incorrectly during a training trial. Token delivery was considered accurate if it occurred only following attending behavior unrelated to the training task (e.g., proper sitting) during baseline or probe trials or following a correct unprompted response during training trials.

For each trial, observers circled “yes” if prompts and tokens were delivered accurately or “no” if prompts and tokens were not delivered accurately. Agreements were defined as trials in which both observers circled “yes” or “no,” and disagreements were defined as trials in which one observer circled “yes” and the other circled “no.” Agreements were divided by agreements plus disagreements, and this ratio was converted to a percentage. The mean percentage of accurate presentation of prompts and tokens was 99.9% across all participants and conditions (range, 98% to 100%). Mean interobserver agreement on accurate presentation of prompts and tokens was 99.9% (range, 98% to 100%).

### *Social Validity*

Nineteen teachers from the behavioral intervention program who were not involved in the study assessed the social validity of the results. Observers watched videotaped segments of each of the participants responding to various vignettes with dolls and puppets. Each segment (approximately 10 to 15 s in duration) showed the instructor presenting a vignette with a doll or puppet and the participant’s response, but did not depict instructor feedback. For all participants, three baseline and three treatment segments were selected randomly from the last three baseline and treatment sessions for the sadness or pain response category and were presented in random order. For Josh, three baseline and three treatment segments were also

presented for the happiness or excitement and frustration response categories. Because virtually no responding occurred during baseline, all baseline segments showed participants responding incorrectly or not at all. All segments from treatment sessions showed responses that had been scored as correct. Observers were given a written questionnaire and for each segment circled “yes” or “no” for the question, “Did the child demonstrate empathy toward the doll or puppet?”

## RESULTS

Figures 1 (Jacob, Luke, Ali) and 2 (Josh) depict the number of motor and vocal empathy responses emitted to displays of affect with training stimuli. During the baseline condition, all participants responded infrequently to displays of affect by the dolls and puppets. Increased pretend-play empathetic responding occurred systematically with the introduction of treatment across participants.

Figures 3 (Jacob, Luke, Ali) and 4 (Josh) depict the number of motor and vocal empathy responses emitted to displays of affect with nontraining probe stimuli. Appropriate responding generalized from training to nontraining vignettes, dolls, and puppets for all participants, increasing systematically with the introduction of treatment with training stimuli across participants. For Josh and Jacob, generalized motor and vocal responses emerged at about the same time. Luke’s generalized motor and vocal responding with nontraining probe stimuli was variable, with vocal responding emerging more quickly than motor responding. For Ali, the opposite was true; generalized motor responding emerged slightly more quickly than generalized vocal responding.

Figures 5 (Jacob, Luke, Ali) and 6 (Josh) depict the number of motor and vocal empathy responses to displays of affect by Person 1 and Person 2 in a nontraining setting. For Josh and Jacob, motor and vocal responding generalized from training dolls and puppets to a teacher

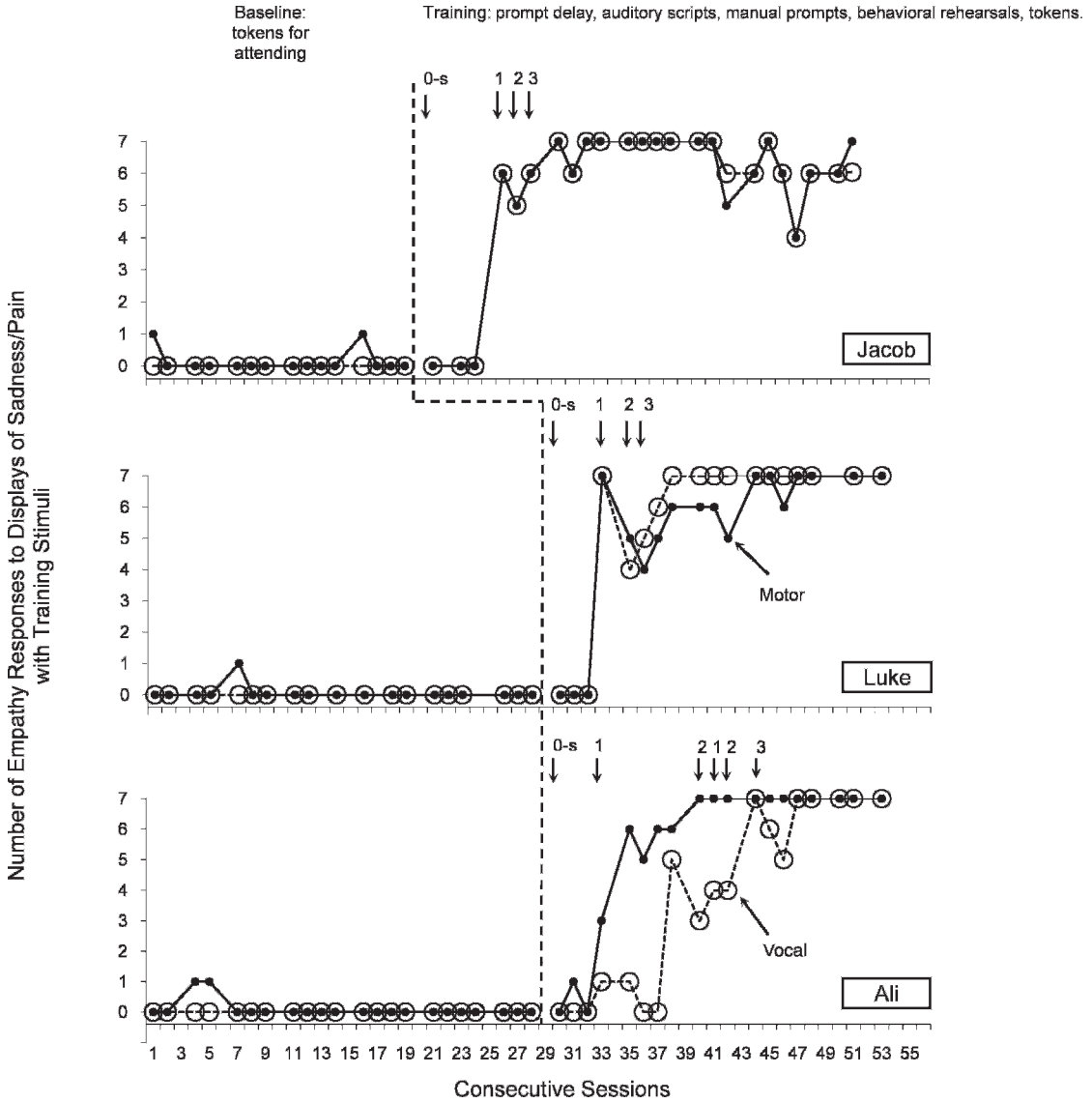


Figure 1. Number of motor (filled circles) and vocal (open circles) empathy responses to displays of sadness or pain with training stimuli for Jacob, Luke, and Ali.

(Person 1) in a nontraining classroom. In addition, both boys responded to displays of affect by another teacher (Person 2) in the nontraining setting, although no baseline measurements were taken with this person (aside from a single observation in Session 25 in the frustration category for Josh). Like Josh and Jacob, Luke and Ali demonstrated generalized motor and vocal empathetic responding with

Person 1 systematically with the introduction of treatment with training stimuli in the training setting, although motor responding during this condition returned to baseline levels during the last two probe sessions. With the brief introduction of training (i.e., vignettes presented by Person 1, 3-s prompt delay, auditory prompts, manual prompts, behavioral rehearsals, tokens delivered for correct responses) in the nontrain-



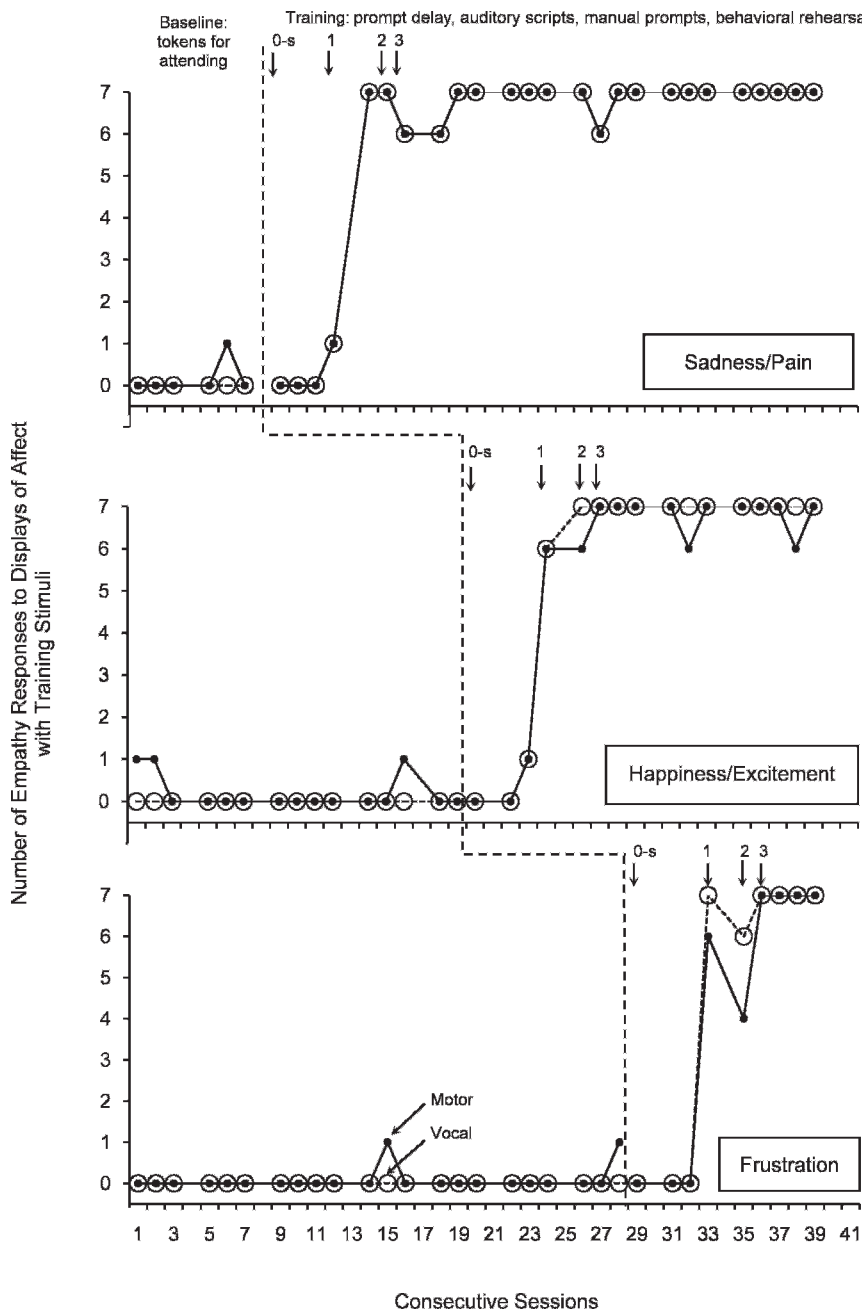


Figure 2. Number of motor (filled circles) and vocal (open circles) empathy responses to displays of affect with training stimuli for Josh.

ing setting (Sessions 55 and 56), motor responding increased to the same high level as vocal responding for both Luke and Ali. Increases in generalized empathetic responding

to Person 2 were also observed for both Luke and Ali, although, as with Person 1, vocal responding was emitted at higher levels than motor responding.

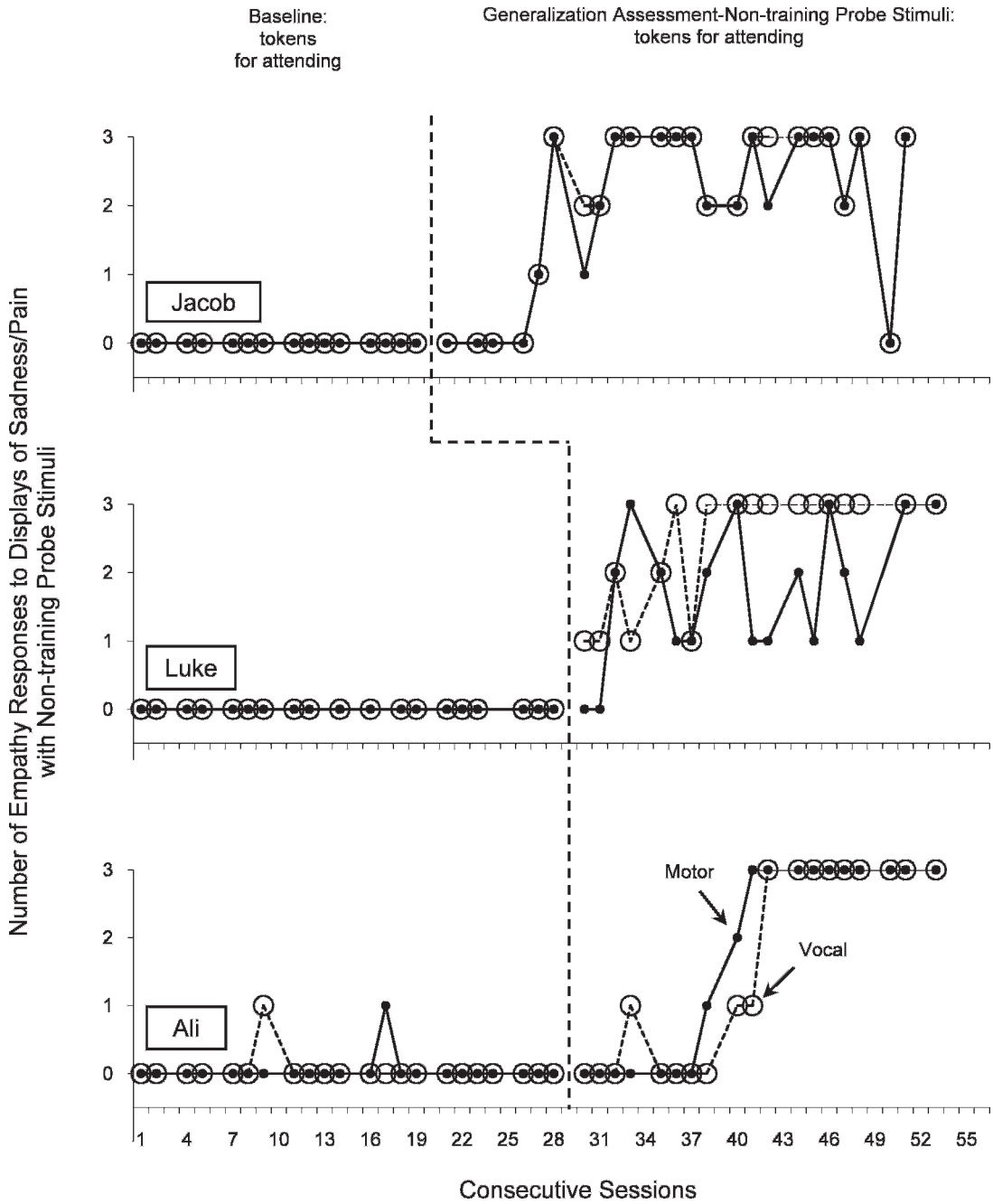


Figure 3. Number of motor (filled circles) and vocal (open circles) empathy responses to displays of sadness or pain with nontraining probe stimuli for Jacob, Luke, and Ali.

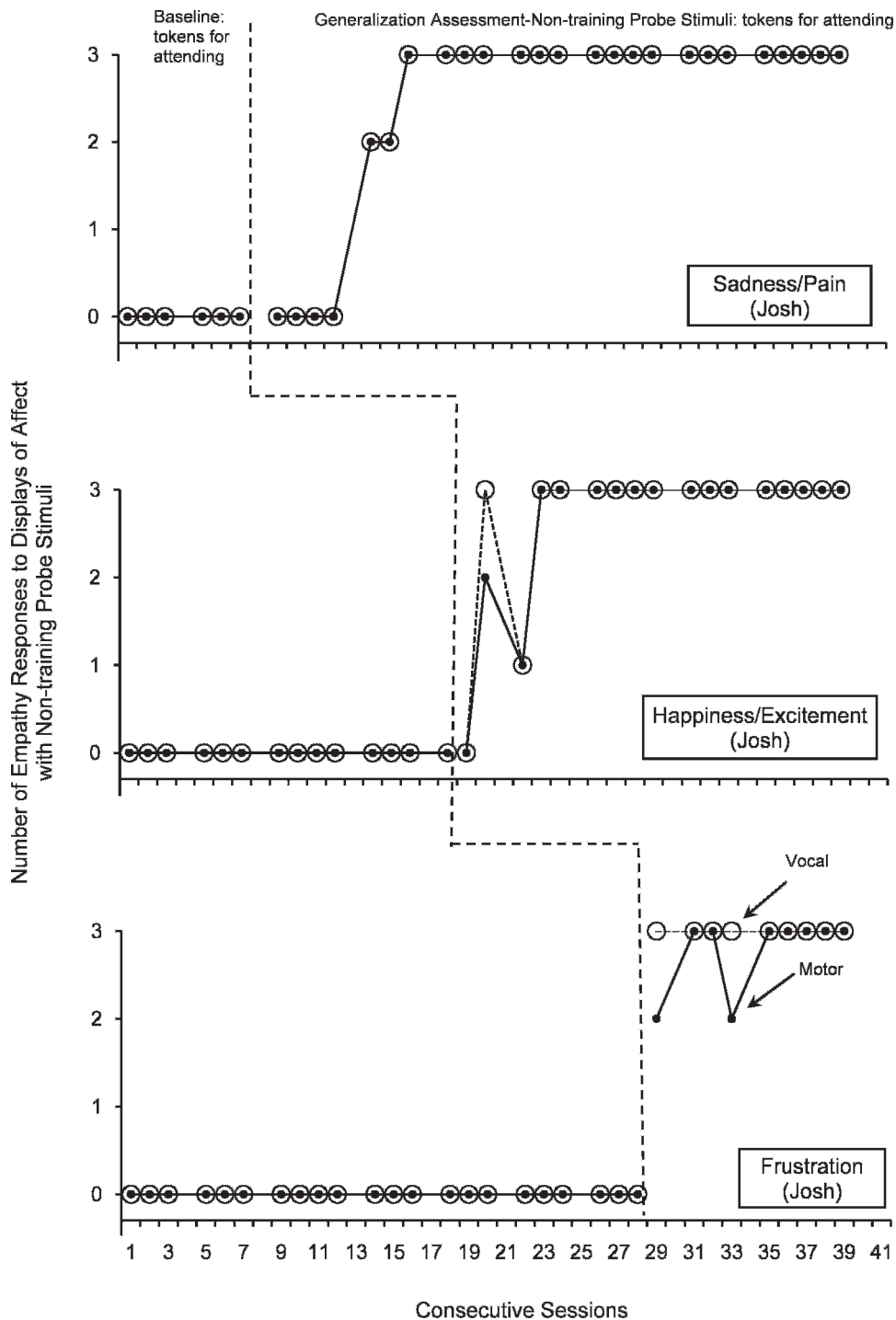


Figure 4. Number of motor (filled circles) and vocal (open circles) empathy responses to displays of affect with nontraining probe stimuli for Josh.

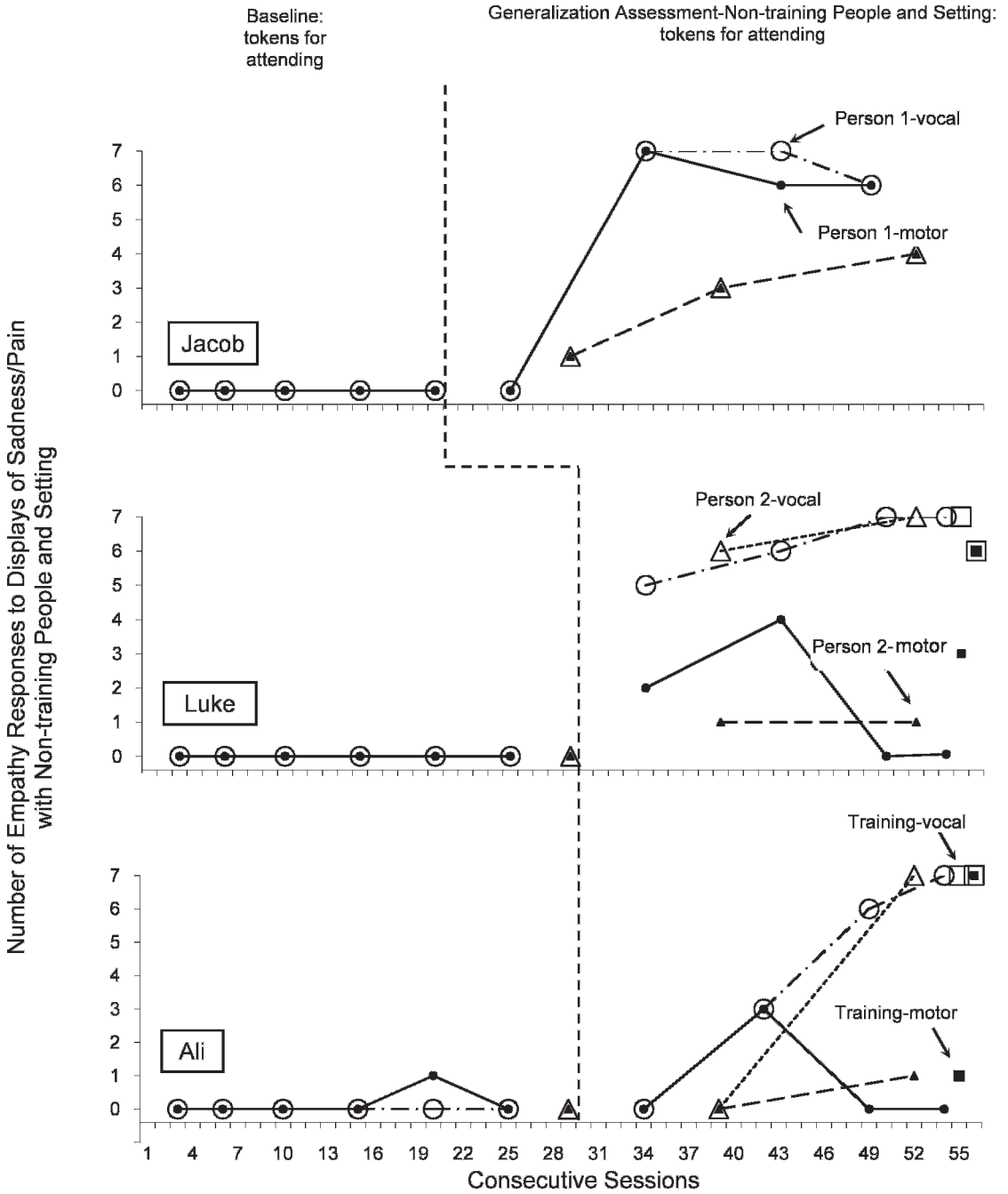


Figure 5. Number of motor (solid lines) and vocal (dashed lines) empathy responses to displays of sadness or pain with Person 1 (open and filled circles) and Person 2 (open and filled triangles) in a nontraining setting for Jacob, Luke, and Ali.

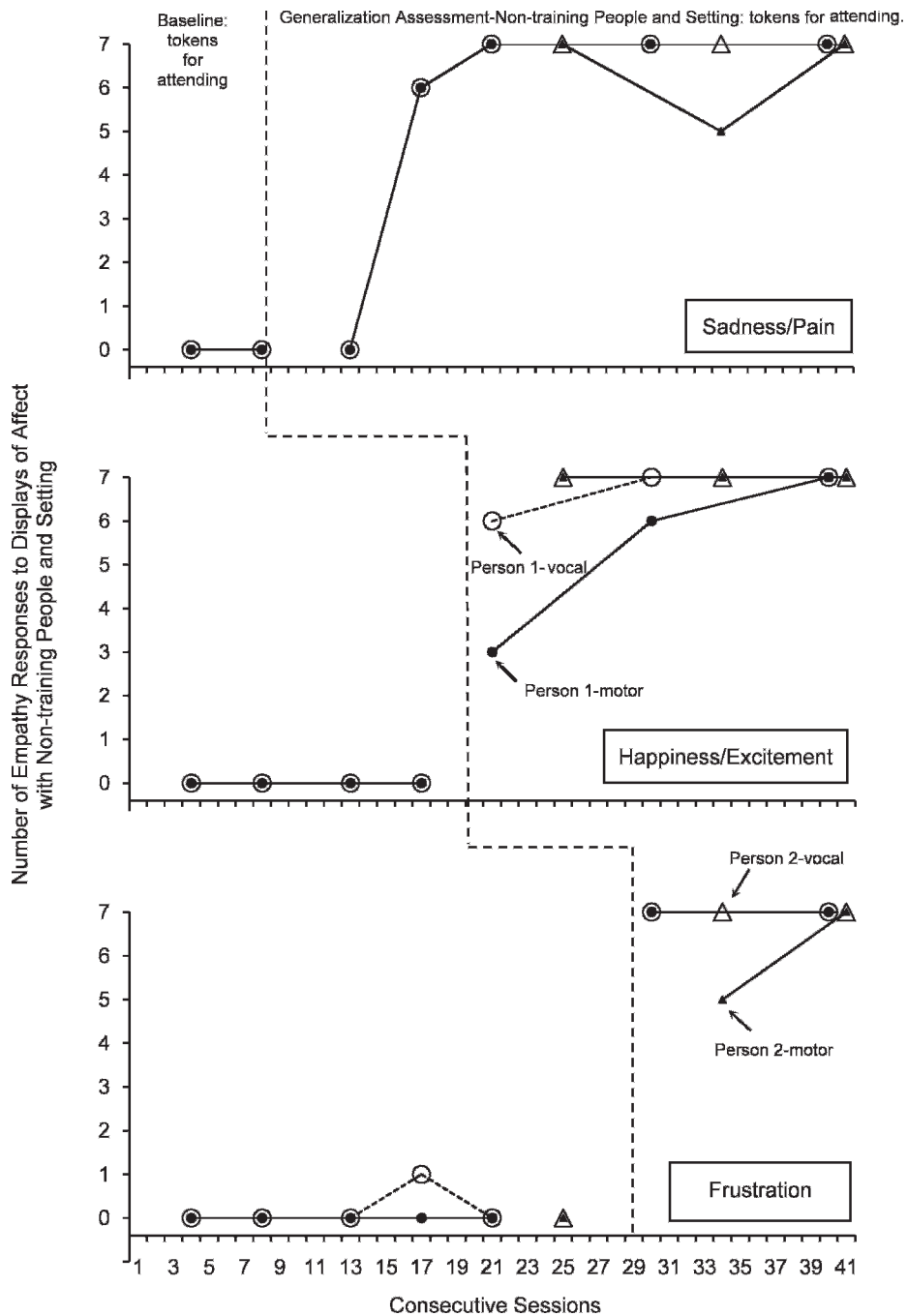


Figure 6. Number of motor (solid lines) and vocal (dashed lines) empathy responses to displays of affect by Person 1 (open and filled circles) and Person 2 (open and filled triangles) in a nontraining setting for Josh.

### *Social Validity*

Data were summarized as the number of baseline and treatment segments scored for empathy (range, 0 to 3). Observers consistently scored empathy in more treatment than baseline segments for all participants and response categories. The mean number of sadness or pain segments scored for empathy was 0.02 (range, 0 to 1) for Jacob and Luke, 0.03 (range, 0 to 1) for Ali, and 0 for Josh during baseline, and 3 for all participants during treatment. For Josh, the mean number of happiness or excitement segments scored for empathy was 0.05 (range, 0 to 1) during baseline and 3 during treatment, and the mean number of frustration segments scored for empathy was 0 during baseline and 3 during treatment.

## DISCUSSION

Because empathetic responding to displays of sadness or pain increased systematically with the introduction of treatment across all participants, we can conclude that the treatment was effective for this group of students. Similarly, because Josh's empathetic responding to displays of affect in three response categories increased systematically with the introduction of treatment, we can conclude that the treatment was effective for Josh. We can also conclude that the treatment was effective in promoting generalized empathetic responding. Specifically, empathetic responding generalized from training to nontraining probe stimuli, increasing systematically with the introduction of treatment with the training stimuli across all 4 participants and all three response categories. In addition, empathetic responding to actual people in a nontraining setting increased systematically with the introduction of treatment with the training stimuli in the training setting for 2 participants. Finally, a group of teachers scored empathy more frequently during videotaped treatment sessions than during baseline sessions. Collectively, our results indicate that children with autism as young as 4 years old can learn to

demonstrate socially relevant empathy skills in pretend-play settings, and that these skills can generalize to pretend-play vignettes and toys not targeted during training. It is important to note that these skills can also generalize to interactions with real people in nontraining settings, although generalization from pretend play to real life may be limited.

Dolls and puppets (rather than actual people) were used for this study because typical peers and other people were not readily available for teaching in this setting, as is frequently the case when teaching students with autism. Teaching through pretend-play vignettes allowed us to address empathy skills without involving other people and allowed easy repetition of trials when needed. Because we did not conduct a component analysis to separate the effects of individual treatment package components, we cannot say how much influence the pretend-play component had over responding. Nevertheless, given the limited generalization to actual people observed for both Luke and Ali, teaching with actual people (e.g., peers, siblings) would likely be a more effective strategy to promote generalization from pretend play to real life. In fact, when treatment was introduced briefly with an actual person in the nontraining setting, both Luke and Ali quickly emitted the target responses. This procedure would more closely resemble those of previous studies (e.g., Gena et al., 1996; Reeve et al., 2007) that demonstrated the effectiveness of components of our treatment package (e.g., modeling, prompting, reinforcement) in teaching other complex social skills.

The types of responses displayed by the participants during baseline supported our hypothesis that deficits in empathy and perspective taking may not indicate the absence of necessary responses but rather an inability to distinguish the conditions under which a specific response (e.g., offering assistance, demonstrating interest) would be appropriate. During baseline, the few responses displayed

were often inappropriate for the given vignette category but appropriate for another (e.g., laughing when the puppet was sad). Therefore, it may be most effective to teach responding separately for different response categories, as occurred in the present study.

Multiple-exemplar training (Stokes & Baer, 1977), in addition to training with complex stimulus compounds for each response category, might have been responsible for the generalization observed to nontraining probe stimuli, people, and settings, because the children learned to attend and respond to a variety of motor and vocal discriminative stimuli for each response category. One can speculate about whether the responses demonstrated by the participants in this study were general enough to be considered empathetic. We presented multiple stimulus compounds in training for the sadness or pain category, and for Josh, we taught empathy responses across a variety of response categories but did not specifically investigate how many exemplars would be required to produce a generalized empathy repertoire. It seems likely that this number differs across individuals, but that training with multiple exemplars should eventually produce a class of responses one could reasonably label as empathetic.

Future researchers may also consider presenting motor and vocal elements of the stimulus compounds separately to determine whether empathetic responding would occur to one element without the presence of the other. We made the decision to present both elements together in an attempt to present the most salient discriminative stimuli possible to our beginning empathy responders. Certainly a more sophisticated generalized empathy repertoire would include responses to motor or vocal displays of affect in isolation.

Differences in generalized responding to Person 1 and Person 2 observed for several of the participants in the nontraining setting may have been due to familiarity with each of the

teachers. Person 1 acted as the prompter in the training setting (but never presented vignettes), whereas Person 2 was a completely unfamiliar teacher. Introduction of training with Person 2 in the nontraining setting, as occurred with Person 1, may have resulted in similar increases in empathetic responding, but this training was not conducted because of time limitations.

One limitation of the study was the cumbersome nature of the treatment package. Language Master scripts were used because they did not require the participants to have reading skills or to interact with the prompter (i.e., when given vocal prompts) before responding to the dolls and puppets. Language Master scripts are also useful because of the ease with which they can be faded. In this case, however, given all of the other treatment components involved, using vocal prompts may have simplified the procedure. This change, in addition to teaching with actual people instead of dolls and puppets might produce a more manageable treatment package.

An additional limitation of this study is the nature of the raters in the social validity assessment. The raters, although not directly involved with the study, were teachers at the behavioral intervention program and so had familiarity with a variety of individuals with autism. Their knowledge of autism and behavioral techniques used to teach complex social skills may have affected their ratings of empathy shown in the videotaped segments. A group of raters inexperienced with autism might have produced a more convincing demonstration of social validity. Future studies may also compare ratings of empathy demonstrated toward puppets and dolls to ratings of empathy demonstrated toward actual people. It is plausible that responses deemed empathetic in a pretend-play situation may not appear as genuine in interactions with people. It would also be informative to assess parent and caregiver ratings of empathy demonstrated in unstructured home and community settings.

Empathy is a socially relevant skill for children with autism because it enhances their ability to engage in other critical prosocial behaviors (Rheingold & Hay, 1980). Children with autism who demonstrate appropriate concern and interest when others are sad, excited, or frustrated will likely increase their future interaction with peers and family members. This may be especially important for children like Josh who are making the transition from special education classrooms to typical classrooms. The results of this study support the conclusions of previous research that behavior-analytic techniques are effective in improving complex social skills such as empathy for children with autism. This seems an important step toward ameliorating an often noted but rarely treated deficit in children with autism.

#### REFERENCES

- Bailey, S. L. (1981). Stimulus overselectivity in learning disabled children. *Journal of Applied Behavior Analysis, 14*, 239–248.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition, 21*, 37–46.
- Bauminger, N., & Kasari, C. (1999). Brief report: Theory of mind in high-functioning children with autism. *Journal of Autism and Developmental Disorders, 29*, 81–86.
- Charlop, M. H., Schreibman, L., & Thibodeau, M. G. (1985). Increasing spontaneous verbal responding in autistic children using a time delay procedure. *Journal of Applied Behavior Analysis, 18*, 155–166.
- Eisenberg, N. (1992). *The caring child*. Cambridge, MA: Harvard University Press.
- Gena, A., Krantz, P. J., McClannahan, L. E., & Poulson, C. L. (1996). Training and generalization of affective behavior displayed by youth with autism. *Journal of Applied Behavior Analysis, 29*, 291–304.
- Halle, J. W., Marshall, A. M., & Spradlin, J. E. (1979). Time delay: A technique to increase language use and facilitate generalization in retarded children. *Journal of Applied Behavior Analysis, 12*, 431–439.
- Harris, S. L., Handleman, J. S., & Alessandri, M. (1990). Teaching youths with autism to offer assistance. *Journal of Applied Behavior Analysis, 23*, 297–305.
- Lovaas, O. I., Koegel, R. L., & Schreibman, L. (1979). Stimulus overselectivity in autism: A review of research. *Psychological Bulletin, 86*, 1236–1254.
- MacNamara, J., Baker, E., & Olson, C. L. (1976). Four-year-olds’ understanding of pretend, forget, and know: Evidence for propositional operations. *Child Development, 47*, 62–70.
- Reeve, S. A., Reeve, K. F., Townsend, D. B., & Poulson, C. L. (2007). Establishing a generalized repertoire of helping behavior in children with autism. *Journal of Applied Behavior Analysis, 40*, 123–136.
- Rheingold, H. L., & Hay, D. F. (1980). Prosocial behavior of the very young. In G. S. Stent (Ed.), *Morality as a biological phenomenon* (pp. 93–108). Berkeley: University of California Press.
- Rheingold, H. L., Hay, D. F., & West, M. J. (1976). Sharing in the second year of life. *Child Development, 47*, 1148–1158.
- Rutter, M. (1978). Diagnosis and definition of childhood autism. In M. Rutter & E. Schopler (Eds.), *Autism: A reappraisal of concepts and treatment* (pp. 139–161). New York: Plenum.
- Schreibman, L., Koegel, R. L., & Craig, M. S. (1977). Reducing stimulus overselectivity in autistic children. *Journal of Abnormal Child Psychology, 5*, 425–436.
- Sigman, M. D., Kasari, C., Kwon, J. H., & Yirmiya, N. (1992). Responses to the negative emotions of others by autistic, mentally retarded, and normal children. *Child Development, 63*, 796–807.
- Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*, 349–367.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children’s understanding of deception. *Cognition, 13*, 103–128.
- Yirmiya, N., Sigman, M. D., Kasari, C., & Mundy, P. (1992). Empathy and cognition in high-functioning children with autism. *Child Development, 63*, 150–160.

*Received October 3, 2006*

*Final acceptance September 4, 2007*

*Action Editor, James Carr*