Using scripts and script-fading procedures to promote bids for joint attention by young children with autism

Joyce L. MacDuff *, Regina Ledo, Lynn E. McClannahan, Patricia J. Krantz

Princeton Child Development Institute, 300 Cold Soil Road, Princeton, NJ 08540, United States
Received 21 November 2006; accepted 24 November 2006

Abstract

A multiple-probe design across participants assessed the efficacy of using audiotaped scripts to promote the joint attention responses of young children with autism. A one-word script ("See") was presented on button-activated voice recorders; recorders were attached to toys and photographs that were placed in areas of the school not typically used for instructional activities. Children were taught to activate the recorders, point to pictures and toys, orient toward a conversation partner, and say the script. After they reliably made these responses, scripts were removed from the recorders and then recorders were removed. All three children learned to make bids for joint attention and after the scripts were faded, bids for joint attention maintained and generalized to untrained materials and to non-training settings.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Joint attention; Script-fading

1. Introduction

Some of the earliest social and communicative behavior to emerge in infants are joint attention responses, which have been defined as the "... use of gestures and eye contact to coordinate attention with another person in order to share the experience of an interesting object or event" (Mundy, Sigman, & Kasari, 1994, p. 389). Two forms of joint attention have been described in the developmental literature—responding to another person's initiations and initiating to others.

* Corresponding author. Tel.: +1 609 924 6280; fax: +1 609 924 4119.
E-mail address: info@pdi.org (J.L. MacDuff).

1750-9467/ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.rasd.2006.11.003
According to developmental and cognitive psychologists, the function of either form is not to obtain an object or activity, but to engage in social interaction. Only recently, have behavior analysts begun to describe the events that may evoke and maintain these different response classes (Dube, MacDonald, Mansfield, Holcomb, & Ahearn, 2004).

In typically developing children, joint attention repertoires begin as early as 9 months, when babies display gaze monitoring (following the gaze of a communication partner), and are fully developed at 16–18 months (Bakeman & Adamson, 1984). These responses are not static, but follow a developmental course; as children develop, shifting of eye gaze begins to be accompanied by gestures, such as holding up and showing objects. Then pointing replaces showing and vocal responses that initially accompany pointing may later replace pointing altogether (Leung & Rheingold, 1981; Rheingold, 1977; Rheingold, Hay, & West, 1976; Tomasello, 1995). These skills are considered to play a critical role in language and social development because of the consequences that typically follow bids for joint attention, namely, the attention of other persons and associated responses, such as smiles, praise, and language models (Jones & Carr, 2004; Mundy & Crowson, 1997; Mundy, Sigman, & Kasari, 1990; Tomasello, 1995).

As typically developing children acquire language, they use it to garner others' attention. But children with autism often do not have enough language to describe events, do not respond to cues to display language in appropriate contexts, and do not draw others' attention to objects and events. Many children with autism who acquire complex verbal repertoires do not spontaneously initiate interaction, or interact only to obtain preferred stimuli (Krantz & McClannahan, 1998).

It is well established that in comparison to typically developing peers as well as to intellectually disabled peers, the behavioral repertoires of children with autism are characterized by deficits in joint attention. Deficits in responding to and initiating bids for joint attention have been documented (Charman et al., 1997; Mundy, Sigman, Ungerer, & Sherman, 1986; Osterling & Dawson, 1994). Indeed, screening tools, such as the Checklist for Autism in Toddlers (Baron-Cohen et al., 1996), include items that assess joint attention skills, and such skill deficits are one of the criteria for diagnosing autism (American Psychiatric Association, 2000).

Several studies have used joint attention skills as outcome measures of interventions that teach children with autism social skills, language, and play (e.g., Mundy et al., 1990), but only a small body of research documents effective strategies that specifically target these skills. Klinger and Dawson (1992) taught a 5-year-old verbal youngster to alternate gaze to initiate joint attention and to respond to joint attention bids during interactions in which interesting and novel activities were presented. Hwang and Hughes (2000) taught non-verbal children with autism, age 2–4 years, to initiate bids for joint attention by alternating gaze, pointing, and showing. However, when these skills were assessed across settings and people, limited generalization was reported. Whalen and Schreibman (2003) taught five young children with autism to initiate and respond to joint attention bids by pointing to objects and shifting gaze. Their treatment package included teaching during child-selected activities, turn taking, manual and gestural prompts, and contingent reinforcement (access to the object) following prompted and unprompted target responses. Generalization of these skills was observed across people and in unstructured conditions for some of the children, but gains did not maintain at treatment levels during follow up.

These studies indicate that further research is required to evaluate procedures for teaching verbal and non-verbal children with autism to respond to and initiate bids for joint attention. Further examination of strategies that enhance maintenance and generalization of these skills is also needed.
In this study, we examined scripts and script-fading procedures as a strategy for teaching joint attention skills to youngsters with autism. These procedures have been documented as effective in teaching conversational skills to learners with autism (Krantz & McClannahan, 1993; Krantz & McClannahan, 1998; Sarokoff, Taylor, & Poulson, 2001; Stevenson, Krantz, & McClannahan, 2000). Written scripts have been used for readers and scripts recorded on magnetic cards or button-activated voice recorders have been helpful to non-readers. When such scripts were systematically faded, children and youths learned to approach and orient to conversation partners, initiate interaction, and respond to peers’ and adults’ social initiations. Recently, we began to examine scripts and script-fading to teach youngsters with autism to make bids for joint attention.

2. Method

2.1. Participants

The three participants – Josh, Hank, and Joe – attended the preschool program of the Princeton Child Development Institute for 5.5 h per day, 5 days per week. All three boys met the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition criteria for autism (American Psychiatric Association, 2000) and had diagnoses of autism conferred by one or more outside agencies. Informed parental consent to participate in the study was obtained for each boy.

Josh, age 4, had been enrolled in the preschool for 0.6 years when the study began. Upon program entry, he did not have any expressive language and displayed tantrum behavior, food refusal, and trichotillomania. Tests administered just prior to the study produced age-equivalent scores of 1–9 on the Preschool Language Scale, and 1–10 on the Vineland Adaptive Composite Scales.

Hank, age 3, received intervention services from the Institute for 0.4 years before the study began. At enrollment, he did not display verbal- or motor-imitation skills, had an extremely limited receptive language repertoire, and engaged in tantrums and self-injurious behavior (head hitting). In the month preceding the study, formal assessment yielded age-equivalent scores of 1–5 on the Preschool Language Scale, and 1–5 on the Vineland Adaptive Composite Scales.

Joe, age 5, entered the program 2 years earlier, with a limited verbal-imitation repertoire; he had matching skills, but no speech. He displayed many stereotypic responses and frequent tantrums. Pre-investigation assessment yielded a Preschool Language Scale age-equivalent score of 1–9; administration of the Vineland Adaptive Composite Scales resulted in an age-equivalent score of 1–10.

Before the study began, all three children learned to imitate a few one-syllable words and to make pointing responses, which they used to mand. They were not observed to use pointing or verbal responses to initiate bids for joint attention.

2.2. Settings

Teaching sessions were conducted in two hallways that the boys infrequently used. Generalization sessions were conducted in a large conference room and in a third seldom-used hallway. The order of use of these settings was randomized, as were the directions in which children traversed the areas. In all four settings, they walked 16.8 m though corridors that were 1.7–2.7 m wide. Two- and three-dimensional stimuli were rotated across locations in these corridors. Stimuli were selected on the basis that they were age-appropriate and of potential interest to young children. They consisted of familiar objects (e.g., telephone, broom); toys (e.g.,
Slinky® drum; stuffed animals (e.g., teddy bear, lion); and character dolls (e.g., Elmo®, Bob the Builder®, etc.). Forty two-dimensional stimuli (pictures and photographs) and 40 three-dimensional stimuli were used during teaching sessions; 10 two-dimensional stimuli and 10 three-dimensional stimuli were used during generalization sessions. The stimuli and locations used in each session were selected using a table of random numbers.

In each session, three two-dimensional and three three-dimensional stimuli were placed in each of two teaching or two generalization settings. These 12 stimuli were displayed on walls, floors, window sills, chairs, and on 59 cm x 43 cm x 30.5 cm clear storage bins. Pictures and photographs were displayed in 20.3 cm x 25.4 cm (8 in. x 10 in.) clear plastic frames that were attached to wooden easels. Small (4 cm in diameter) circular button-activated recorders (Mini-Me® voice recorders available from Voice Express), were attached to two- and three-dimensional objects and played the previously recorded script (“See”) when a child pressed the button in the center of the device.

2.3. Pre-investigation teaching and assessment

Prior to the study, the youngsters' verbal imitation skills were assessed; the instructor modeled the word “See,” and one-syllable words never used during teaching were modeled on a button-activated recorder (e.g., “me,” “hat”). Pointing skills were tested using a 21.5 cm x 27.9 cm (8.5 in. x 11 in.) book in which one sticker was placed on each page. The book was held just out of a child’s reach and he was asked to “Point to the sticker.” Verbal prompts, models, and token rewards were used to teach verbal imitation; graduated guidance and token rewards were used to teach pointing responses. After all three boys correctly completed 10 verbal imitation and 10 pointing tasks in a single session, baseline began.

2.4. Dependent variables

A verbal bid for joint attention was defined as a single word or phrase, or an approximation to a word or phrase that pertained to an object in the immediate environment. Verbal responses were scored if they were contextual (e.g., saying “ball” in the presence of a ball, but not in the presence of a puzzle) and occurred in the absence of stereotypic or disruptive behavior. Responses were scored only if the child oriented toward the interaction partner, at least momentarily, within 2 s before or after a verbal response.

Scripted bids for joint attention were those that matched a script recorded on a button-activated voice recorder. The script used in this study was “See.”

Unscripted bids for joint attention were words or phrases said in the absence of the script, or verbal productions that differed from the script by the addition of a word or words. For example, “See horse” was scored as unscripted because the script was “See.”

Pointing was defined as isolating the index finger of either hand and directing it toward a stimulus. Pointing was scored whether or not it was paired with a verbal response, but only if the child oriented toward his interaction partner, at least momentarily, within 2 s before or after pointing.

2.5. Experimental conditions

Verbal bids for joint attention (scripted and unscripted) and pointing to bid for joint attention were measured during baseline, teaching, generalization, and maintenance in a multiple-probe
design across the three boys. In each condition, the instructor who was the prompter and interaction partner took a child to the end of a hallway or to the entrance of the conference room and said, "Let's walk this way," or "Let's go over here." The instructor followed the child, remaining within 1 m of him while walking in the corridor. After each bid for joint attention, the instructor responded with a brief comment that included the name of the stimulus (e.g., "Cars go fast" or "Pretty butterfly"). During each session, a youngster walked through two corridors.

2.5.1. Baseline

During this condition, no prompts were delivered. The instructor carried the youngster’s token board and tokens, but tokens were delivered only for appropriate walking in the absence of stereotypy and disruptive behavior; a maximum of one token was delivered in each corridor (a maximum of two tokens per session).

2.5.2. Teaching

Initially, button-activated recorders were placed on all of the pictures and toys and if a child attempted to walk past any of the stimuli without making a bid for joint attention, he was manually guided to press the button to activate the voice recorder. Manual guidance was also used if he did not isolate an index finger when activating a recorder. If he did not say the script ("See"), or did not orient toward her while saying the script, the instructor guided him to use his index finger to again activate the recorder. Tokens were delivered following prompted bids for joint attention. Special snacks (i.e., snacks that were not available at any other time during the school day) were dropped in a cup each time a boy made an unprompted bid for joint attention.

After the youngsters independently made bids for joint attention when encountering at least 11 of the 12 stimuli, script-fading began. The first fading step consisted of erasing the audiotaped model ("See") from the button-activated recorder, so that only blank recorders were attached to the stimuli. In the second fading step, recorders were removed from all stimuli. Concurrent with the second fading step, special snacks were no longer provided. After Josh and Hank had at least nine bids for joint attention in each of five consecutive sessions, the number of sessions was reduced from two sessions a day to one session a day, then to two sessions every 3 days, and then to one session every 2 days.

Joe’s bids for joint attention decreased after the recorders were removed, and therefore a more gradual fading sequence was implemented for him. Five sessions after the recorders were removed (at Session 69), blank recorders were replaced on all stimuli, where they remained for one session. At Session 70, half of the recorders were removed but recorders remained on the first, third, and fifth stimulus in each corridor. After seven additional sessions (at Session 77), blank recorders were removed from Stimuli 1, 3, and 5 and were placed on Stimuli 2, 4, and 6 in each corridor. On Session 78, the number of blank recorders was reduced to two (on Stimuli 1 and 3) and on Session 79 there was one blank recorder (on Stimulus 1) in each location. At Session 80, after 11 sessions, the blank recorders were completely faded, and at Session 86, after five consecutive data points on which there were seven or more bids for joint attention, the number of sessions conducted was reduced from two per day to one per day.

2.5.3. Generalization probes

Generalization was assessed in locations never used during teaching (a conference room in which tables were arranged to create a corridor, and a different hallway). The stimuli used during
generalization probes were never used during teaching and no prompts were delivered. The instructor carried the youngster’s token board and tokens were delivered for appropriate walking. Only one token was delivered in each corridor (a maximum of two tokens per session).

2.5.4. Maintenance

During this condition, which was identical to baseline, no prompts were delivered. The instructor carried the youngster’s token board and tokens were delivered for appropriate walking. A maximum of one token was delivered in each corridor (i.e., a maximum of two tokens was delivered during each session).

2.6. Measurement procedures

Observation began when the instructor gave an invitation, such as “Let’s go over here” and continued until the child walked 16.8 m through each of two corridors or until 3 min elapsed in each corridor. Observers followed the youngster and the instructor through the corridors, remaining approximately 1 m behind the instructor and 1 m distant from one another.

A continuous event recording system was used to record pointing and unscripted bids for joint attention during the observation period. Scripted bids for joint attention were scored using a per opportunity measure. An opportunity was defined as encountering one of six teaching or generalization stimuli while walking through a corridor; the next opportunity occurred when the youngster approached another stimulus. Thus, only one scripted response was scored if, upon approaching a toy car, a child repeated the script “See” twice while orienting to his partner. However, if he said the script “See” while near the car, approached and said the script “See” while near a picture of a clown, and then returned to the toy car and said “See,” three scripted responses were scored.

2.7. Interobserver agreement

Percentage agreement (means and ranges) on scripted and unscripted bids for joint attention and pointing responses is shown in Table 1. Percentage agreement on scripted bids for joint attention was calculated by dividing total number of agreements by total number of agreements plus disagreements, and multiplying by 100. Percentage agreement on unscripted bids for joint attention and pointing was calculated by dividing the smaller total by the larger total and multiplying by 100. Interobserver agreement was obtained on 80% of baseline and generalization sessions, 36% of teaching sessions, and 40% of maintenance sessions.

<table>
<thead>
<tr>
<th>Child</th>
<th>Scripted bids</th>
<th>Unscripted bids</th>
<th>Pointing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Range</td>
<td>Mean</td>
</tr>
<tr>
<td>Josh</td>
<td>96</td>
<td>83–100</td>
<td>98</td>
</tr>
<tr>
<td>Hank</td>
<td>94</td>
<td>83–100</td>
<td>99</td>
</tr>
<tr>
<td>Joe</td>
<td>97</td>
<td>93–100</td>
<td>99</td>
</tr>
</tbody>
</table>
3. Results

Fig. 1 shows the three participants’ scripted and unscripted bids for joint attention across conditions. None of the children initiated bids for joint attention during baseline, with the single exception that Hank said one unscripted word ("Elmo") on a generalization probe.

By the second teaching session, all three boys began to say the script. In 12–19 sessions, they met criterion for saying the script “See” without prompts, after which fading began. Fading steps 1 and 2 (erasing “See” from the recorders and removing the recorders), and discontinuing the delivery of special snacks were completed for Josh and Hank in nine and five sessions, respectively. This sequence for Joe, which included more gradual removal of blank recorders, was completed in 18 sessions. After fading steps were completed and snacks were removed, mean number of unscripted bids for joint attention was 13 for Josh (range = 9–21), 10 for Hank (range = 6–12), and 11 for Joe (range = 7–15).

During maintenance, when prompts and scripts were absent, unscripted bids for joint attention remained relatively unchanged from teaching. Mean number of unscripted bids for joint attention during maintenance was 13 for Josh, 11 for Hank, and 11 for Joe.

Generalization probes during teaching and maintenance showed that bids for joint attention generalized to materials not used during teaching and to environments in which teaching did not.
Table 2
Number of pointing bids for joint attention across conditions

<table>
<thead>
<tr>
<th>Child</th>
<th>Baseline Mean</th>
<th>Range</th>
<th>Teaching Mean</th>
<th>Range</th>
<th>Maintenance Mean</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teaching sessions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Josh</td>
<td>0</td>
<td></td>
<td>7</td>
<td>1-12</td>
<td>9</td>
<td>7-10</td>
</tr>
<tr>
<td>Hank</td>
<td>0</td>
<td></td>
<td>5</td>
<td>1-10</td>
<td>3</td>
<td>1-5</td>
</tr>
<tr>
<td>Joe</td>
<td>0</td>
<td></td>
<td>4</td>
<td>0-9</td>
<td>3</td>
<td>1-8</td>
</tr>
<tr>
<td>Generalization probes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Josh</td>
<td>0</td>
<td></td>
<td>6</td>
<td>2-9</td>
<td>7</td>
<td>5-8</td>
</tr>
<tr>
<td>Hank</td>
<td>0</td>
<td></td>
<td>2</td>
<td>0-5</td>
<td>4</td>
<td>3-4</td>
</tr>
<tr>
<td>Joe</td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>5</td>
<td>3-7</td>
</tr>
</tbody>
</table>

occur. Josh made 7–15 bids (mean = 11); Hank made 6–12 bids (mean = 9); and Joe made 6–9 bids (mean = 7).

Table 2 presents the pointing responses displayed by the three participants in each condition and during generalization probes. In baseline, none of the youngsters pointed to initiate a bid for joint attention.

After teaching began, pointing occurred because it was required in order to activate the voice recorders, but after the second fading step, when the recorders were removed, the number of pointing responses decreased slightly for all participants. On generalization probes conducted during teaching, Josh’s pointing remained relatively unchanged, there was a notable decrease in Hank’s pointing, and Joe’s pointing responses did not transfer across materials and settings. On generalization assessments during maintenance, although unscripted bids for joint attention remained relatively stable (see Fig. 1), Josh’s pointing decreased slightly and Hank’s and Joe’s pointing increased in comparison to probes conducted during teaching.

4. Discussion

In previous investigations (Krantz & McClannahan, 1993; Krantz & McClannahan, 1998; Stevenson et al., 2000), scripts were embedded in photographic activity schedules that were used to cue interaction. In this study, scripts were attached to objects and photographs that were placed throughout school corridors. We endeavored to produce stimulus conditions under which typically developing children would engage in bids for joint attention (that is, when they encountered novel or interesting stimuli). Prior to this study, all three boys had learned to interact with instructors when cues were embedded in activity schedules. In addition, they had learned the labels of many of the stimuli.

Baseline data document the boys’ deficits in joint attention skills. During this condition, only one bid for joint attention was observed; Hank said, “Elmo” when he encountered an Elmo™ plush toy (a highly preferred item) during a generalization probe. Josh scrutinized the toys and pictures when walking past them, but he did not interact with the instructor, although she remained within 1 m of him and consistently visually attended to him. Joe oriented toward many of the items, but typically walked rapidly down corridors and did not make bids for joint attention.

In the first teaching session, all three children said the script “See,” and all three participants met script-fading criteria after a relatively small number of teaching sessions (12–19). The
number of sessions before criteria were met was greatest for Hank (see Fig. 1). In early teaching sessions, he walked past stimuli and was prompted to return to them and to point and activate the recorder.

When scripts were present during teaching, the number of bids for joint attention never exceeded the number of stimuli encountered. When a younger said the script, the instructor commented on the toy or picture, and immediately dropped a snack in a cup. Prompts to engage in further conversation were never delivered and the learners often resumed walking down the corridor. This may be a result of prior discrete-trial training, during which tangible rewards typically immediately follow correct unprompted responses (McClannahan & Krantz, 1997). Dropping the snack in a cup may have served as a cue that the interaction had ended, prompting them to continue to walk down the hallway and repeat the sequence, rather than to engage in ongoing interaction about the item. This indicates that script-fading and discontinuation of tangible rewards was a key aspect of the intervention package.

When voice recorders were removed, the number of bids for joint attention exceeded the number of stimuli presented; in 15 sessions during teaching and maintenance, Josh said more than 12 unscripted bids for joint attention. He also appropriately manipulated and explored toys for longer periods. Joe also made multiple bids for joint attention about certain toys and pictures.

Anecdotal observation indicated that all three boys not only said the script, but also used novel language during bids for joint attention. Often, unscripted language consisted of object labels (e.g., “Bike”). Josh paired the script with labels to form phrases (e.g., “See plane”). Future research might assess whether the instructor’s language models are later used by the participants. This would provide empirical support for the importance of joint attention skills in fostering language development (Jones & Carr, 2004; Tomasello, 1995).

As previously noted, Hank and Joe made fewer pointing responses during maintenance than during the teaching condition, although they continued to make many unscripted bids for joint attention. This result was not unexpected—pointing and showing responses have also been noted to decrease in typically developing children as they acquire language (Leung & Rheingold, 1981; Rheingold, 1977; Rheingold et al., 1976).

In this study, children were taught to initiate bids for joint attention by pointing to objects and saying a script (“See”). All three children made unscripted bids for joint attention about novel stimuli, in settings that were never associated with teaching. Further, their bids for joint attention maintained when voice recorders and tangible rewards were absent and when no prompts were delivered. Script-fading was effective in teaching youngsters with autism to make bids for joint attention, thereby addressing a critical skill deficit. Like typically developing young children, the participants engaged in spontaneous interaction about novel stimuli, providing numerous opportunities for an available adult to teach additional language.

References


