COMPARING THE EFFECTS OF DESCRIPTIVE COMMENTS VERSUS DESCRIPTIVE COMMENTS PLUS PROMPTED TRIALS ON CHILDREN'S LETTER NAMING

By

Kiersten Ann Kinder

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements for

the degree of

DOCTOR OF PHILOSOPHY

In

Special Education

December, 2010

Nashville, Tennessee

Approved:

Dr. Mary Louise Hemmeter

Dr. Ann Kaiser

Dr. Patricia Snyder

Dr. Mark Wolery

ABSTRACT

The purpose of this study was to evaluate the relative effects of a commenting intervention and a commenting plus trials intervention for teaching preschool children letter naming during play activities. For each intervention, descriptive comments were systematically delivered to expose children repeatedly to targeted letter names in a salient context. The commenting plus trials intervention included the use of a constant time delay instructional procedure to deliver prompted trials. An adapted alternating treatments single-subject experimental design was used to compare the effects and efficiency of each intervention on the acquisition, maintenance, and generalization of children's letter naming. In addition, three measures were used to compare the relative efficiency of the interventions. The findings varied across children. Two children acquired, maintained, and generalized letter naming skills under both interventions. For the other two children, only one intervention was effective, and a different intervention was effective for each child. When the interventions were compared in terms of efficiency, differences were minimal. Across participants, the descriptive commenting intervention required slightly fewer sessions but more letter exposures to reach criterion than the descriptive commenting plus trials intervention. Total time to reach criterion, however, was equivalent across the two interventions. Generalization and maintenance data were similar across the two interventions. Implications for practice and future research are discussed.

To my parents, Bob and Shirley Crafton, for first inspiring me to become an educator and supporting me with unconditional love every step of the way.

To my husband, Randy, with gratitude for the many sacrifices you have made to help me achieve this goal. I love you with all my heart and could not have done this without your support.

ACKNOWLEDGEMENTS

This dissertation was supported by the Early Childhood Special Education Leadership Grant funded by the Office of Special Education and a Special Education Endowment Award.

First of all, I would like to acknowledge the overwhelming support of my dissertation advisor, Dr. Mary Louise Hemmeter. I am forever grateful for the many opportunities you have provided for me to learn and grow. Thank you for always believing in me and keeping me focused on my path. Also, I am grateful for the continuing support of my dissertation committee, Drs. Ann Kaiser, Patricia Snyder, and Mark Wolery. Thank you for your guidance on my journey into the world of research - it has been an honor to learn from each of you. Also, I would not have made it here without the influence of three strong women whom I feel fortunate to consider mentors and dear friends. Thank you Sharon Andersen, Donna Nylander, and Michaelene Ostrosky for not only pointing me towards this path, but encouraging me to remain on it along the way.

Many thanks to Jessica Hardy, Adrienne Golden, Elizabeth Will, and Ronnie Benson for their work on this study. I am grateful for the support I continue to receive from many friends and colleagues I met at Vanderbilt University, especially Jessica Hardy, Erin Barton, Amy Casey, Jessie Morris, Alana Schnitz, and the entire “Fab 14” cohort. Without a doubt, my doctoral program experience would not have been as successful, productive, or enjoyable had I not shared it with my dear friend Kathleen Artman Meeker. I do not know what I would have done without you sitting next to me (and then behind me) these last four years. Thank you.

Thanks to my family, church family, and friends who have provided me with much love and prayers as I made this journey. Above all, thanks be to God for His abundant grace and blessings upon my life. Proverbs 3:5.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Since 1991, the number of young children who attend preschool has steadily increased with more than 80% of all 4-year-olds currently attending some form of preschool program (Barnett, Epstein, Friedman, Boyd, & Hustedt, 2008). In recent years, pressure to reduce the achievement gap between America's students and those in comparable countries has resulted in an increased focus on the use of early learning standards and accountability in schools (Copple & Bredekamp, 2009). This, in addition to research demonstrating the benefits of early instruction for literacy and math skills (Clements, Sarama, & DiBiase, 2004; National Institute for Literacy, 2008), has resulted in an increased emphasis on teaching pre-academic skills in preschool settings. A content analysis of state-developed early learning standards revealed a disproportionate emphasis on standards in the language and cognitive domains, particularly early literacy and logic-mathematics skills (Scott-Little, Kagan, & Frelow, 2006). While early learning standards focus on the content of what should be taught, they generally do not offer specific guidelines for how that content should be taught (Grisham-Brown, 2008; Scott-Little, Kagan, & Frelow, 2003). The increased emphasis on teaching pre-academic skills to preschool children raises the question of how and when these skills should be taught.

The Individuals with Disabilities Education Act of 2004 (IDEA, 2004) requires that children with disabilities have access to the general education curriculum. Ensuring access to the general education curriculum requires the use of appropriate strategies for designing and implementing instruction for children with special needs (Nolet & McLaughlin, 2005). This

highlights the need for instructional approaches that are effective for teaching children with special needs while ensuring access to the general education curriculum. In preschool classrooms, accessing the general education curriculum involves participation in typically occurring classroom activities and routines (Grisham-Brown, 2008). Recommended practices in early childhood education and early childhood special education have supported embedding instruction into daily routines and activities (Copple & Bredekamp, 2009; Wolery, 2005). Embedding instruction into typically occurring classroom activities and routines offers a possible means for ensuring all children learn essential skills while being included in the preschool curriculum (Grisham-Brown, Pretti-Frontczak, Hawkins, & Winchell, 2009).

Embedded instruction serves as an umbrella term for the practice of delivering instruction during ongoing activities and routines (Schepis, Reid, Ownbey, & Parsons, 2001; Snell, 2007). The use of this term varies widely in the literature, but may be broadly defined as teaching procedures that are implemented in the context of ongoing routines and activities in the classroom. This type of instruction appears throughout the early childhood special education literature in a variety of approaches including: individualized curriculum sequencing model (Guess et al., 1978; Holvoet, Guess, Mulligan, & Brown, 1980), naturalistic teaching (Warren & Kaiser, 1986; Rule, Losardo, Dinnebeil, Kaiser, & Rowland, 1998), activity-based instruction (Bricker, 1989; Pretti-Frontczak & Bricker, 2004), and transition-based teaching (Werts, Wolery, Holcombe-Ligon, Vassilaros, & Billings, 1992). These approaches have similarities and differences in their procedural components, but all include delivering instruction during typically occurring classroom activities and routines and often include the use of a mand, or an instruction to respond, as part of their instructional sequence.

Another type of intervention that can be embedded into typically occurring classroom activities and routines involves the use of comments without an explicit instruction to respond (Cole & Dale, 1986; Harris & Reichle, 2004; Kaiser & Trent, 2007; Yoder, et al., 1995). While this type of approach has primarily been used to teach social and communication skills, it has been applied in a small number of studies to one type of vocabulary development which is considered one type of pre-academic skill (Girolametto, Weitzman, & Clements-Baartman, 1998; Kouri, 2005; Pemberton & Watkins, 1987; Ellis Weismer, Murray-Branch, & Miller, 1993; Wilcox, Kouri, & Caswell, 1991). Commenting interventions have the potential to address some of the possible concerns associated with interventions that include a mand. Several reasons for not including mands as part of intervention procedures have been discussed in the literature. Continually manding to elicit a response from a child who is unwilling or unable to participate verbally could lead to fewer social interactions and therefore fewer exposures to appropriate verbal models (Rice, 1993; Yoder & Warren, 2001). Delivery of a mand during play could interrupt the child's engagement in an activity or redirect the child's focus of attention during play. When an adult delivers a mand, it sets up an expectation for the child to perform, which could be anxiety provoking (DeThorne, Johnson, Walder, & Mahurin-Smith, 2009). Also, adult-child interactions that include a higher concentration of mands can lead to conversational exchanges that do not reflect the natural patterns of social interaction (Salmon, Rowan, & Mitchell, 1998).

To investigate what is known about these types of embedded interventions, the literature was reviewed and focused on three main topics. First, studies that used mands as a component of the embedded intervention to teach pre-academic skills were reviewed (Kinder, 2009). Second, studies which examined the use of interventions that did not include a verbal mand to teach pre-

academic skills were reviewed. Finally, studies that compared these two types of interventions to teach pre-academic skills were reviewed.

Embedded Instruction that Included the Use of a Mand

Kinder (2009) reviewed the literature on how embedded instruction has been used to teach pre-academic skills. A systematic search of the literature resulted in the identification of 15 single-subject experimental design studies that used embedded instruction to teach at least one pre-academic skill to preschool children (Bambara, Warren, & Kosimar, 1988; Chiara, Schuster, Bell, &Wolery, 1995; Daugherty, Grisham-Brown, & Hemmeter, 2001; Studies 1 & 2 in Fox & Hanline, 1993; Grisham-Brown, Schuster, Hemmeter, & Collins, 2000; Studies 1 & 3 in Horn, Lieber, Li, Sandall, & Schwartz, 2000; Losardo & Bricker, 1994; Studies 1 & 2 in Peck, Killen, & Baumgart, 1989; Werts, et al., 1992; Wolery, Doyle, Gast, Ault, & Simpson, 1993; Wolery, Anthony, & Heckathorn, 1998; Ziolkowski & Goldstein, 2008). To ensure all relevant articles were included, identical search procedures were conducted to identify additional articles published since the review was conducted in May 2009. However, no additional articles were found that met the original inclusion criteria. The inclusion criteria were as follows: (a) the study used single-subject experimental methodology, (b) at least one subject had to be under 6 years of age, (c) the study had to be conducted in a preschool setting in the United States, (d) the independent variable used in the study had to include embedded instruction defined for this review as teaching procedures that were implemented during ongoing routines and activities within the classroom, (e) the independent variable had to be described in enough detail to determine the procedures used for instruction, (f) child data had to be reported as a dependent variable, and (g) at least one pre-academic skill had to be included as a target skill. Examples of

pre-academic skills included matching or sorting; naming shapes, numerals, colors, letters, or objects; counting; demonstrating spatial concepts; or expressively identifying first names, rhyming pairs, or sight words.

The 15 studies identified above included the following single subject designs: 13 used a demonstration design (i.e., multiple-baseline across target skills or participants, multiple-probe across target skills or participants), two used comparison designs (i.e., alternating treatment, adapted alternating treatment), and one used an AB design. Across the 15 studies, there were 56 participants. The overall mean age of the participants was 53.5 months (range: 31-75 months). Information on the socioeconomic status of participants was not provided in most of the studies; when it was provided, the majority of participants were identified as low income (Werts et al., 1992; Wolery et al., 1998; Ziolkowski & Goldstein, 2008). Most participants had some type of identified disability (91%), the others were typically developing (9%). The most commonly identified disability was speech-language impairment (39%). Instruction was typically delivered by a member of the classroom team (87%; i.e., lead teacher, assistant teacher, speech pathologist, or practicum student). In two studies, researchers not associated with the classroom served as the interventionists (Chiara et al., 1995; Ziolkowski & Goldstein, 2008). All 15 studies were implemented in settings described as either childcare or preschool programs, and most programs included children with disabilities (87%).

Across the 15 studies that used embedded instruction, 24 different pre-academic skills were targeted for instruction. Half of the pre-academic target skills were labeling objects or pictures (Chiara et al., 1995; Losardo & Bricker, 1994; Study 2 in Peck et al., 1989; Wolery et al., 1998). Other targeted skills included matching, name recognition, sorting, sight word identification rhyming, and initial sound knowledge. Instruction on pre-academic skills was

delivered during five different types of activities (i.e.; circle time, small group, or free choice play) or routines (i.e.; meal time, self care, or transitions). Transitions were considered a routine because they typically occur at the same time of the day and have the same steps for completion, similar to hand washing or preparing for snack. For example, a transition into the classroom might always include hanging up coats, unpacking backpacks, and choosing a book to read. Across the 15 studies, free choice play was the most frequent activity in which instruction was delivered (47%). Transitions were the most frequent routine in which embedded instruction was delivered (27%). For example, Chiara et al. (1995) delivered instruction on picture naming during whole group transitions (e.g., the class ends large group and lines up at the door) and individual transitions (e.g., a child finishes a puzzle and selects a new game).

Various procedures were used to teach the target skill. Ten studies identified a specific instructional procedure: six used constant time delay (CTD; Chiara et al., 1995; Daugherty et al., 2001; Grisham-Brown et al., 2000; Werts et al., 1992; Wolery et al., 1993; Wolery et al., 1998), two used a naturalistic teaching procedure (Studies 1 & 2 in Fox & Hanline, 1993), one used a least-prompt training strategy (Bambara et al., 1988), and one used a sentence completion strategy (Ziolkowski & Goldstein, 2008). All 10 studies used procedures that included a mand. A mand is a request for the child to verbalize (Rogers-Warren & Warren, 1980). Mands took various forms including task directions, trial cues, and a verbal request for a response (examples provided below). The adult provided various prompts to elicit a child response. For example, Daugherty et al. delivered a task direction (e.g., “Count the crackers”, “Give me seven blocks”) followed by an instruction to imitate a model if the child did not respond. Similar procedures were used in the other five studies that used CTD. The naturalistic teaching procedure used by Fox and Hanline included a least-to-most prompt hierarchy ranging from time delay (e.g.,

pausing with an expectant look) to a model (e.g., "Say red."). Bambara et al. delivered a trial cue (e.g., "Find your name.") and followed with an individualized prompting procedure. Ziolkowski and Goldstein delivered a mand during small group storybook reading by calling a child's name, making a statement about the rhyme or letter sound presented in the story, and pausing for the child to respond, (e.g., "Johnny, bear starts with the /b/ sound. Bear starts with the _____sound"). Regardless of specific procedure used, target skill, or context, in all 10 studies a verbal mand was used to prompt the child to respond.

Five studies did not identify a specific instructional procedure (Studies 1 & 3 in Horn et al., 2000; Losardo & Bricker, 1994; Studies 1 & 2 in Peck et al., 1989); rather, they used global descriptions about how instruction was provided such as using models, giving natural cues, and prompting. In both Horn et al. studies, teachers were given general directions to plan opportunities to deliver instruction on targeted skills throughout the day and support children through natural cues; modeling; and verbal, non-verbal, and physical prompting. Inadequate information was provided to discern how often each support was used during instruction. Losardo and Bricker compared two procedures, activity based intervention and direct instruction. When activity-based intervention was used, a variety of antecedents such as general comments, models, questions, and instructions to respond were used to elicit the target responses. While procedures in the Peck et al. studies were not described specifically, the teacher generated suggestions for how to use prompts during embedded instruction during consultation with the researcher. Prompts were defined as any behavior demonstrated by the teacher in a direct effort to elicit a targeted response from the child. These prompts could include verbal prompts (e.g., directions, modeling), gestural cues (e.g., pointing hand signals, shaking head), or physical guidance. It was not possible to discern exactly which methods were used during intervention.

Despite the lack of specified instructional procedures, the descriptions in all five studies indicate that some type of mand was used.

The majority of studies were coded as having strong or mixed effects on the acquisition of pre-academic skills (68%; Bambara et al., 1988; Chiara, et al., 1995; Daugherty et al., 2001; Studies 1 & 2 in Fox & Hanline, 1993; Losardo & Bricker, 1994; Study 1 in Peck et al., 1989; Werts et al., 1992; Wolery, et al., 1993; Wolery et al., 1998). The remaining five studies received a coding of weak effects on the acquisition of pre-academic skills (Grisham-Brown et al., 2000; Studies 1 & 3 in Horn et al., 2000; Study 2 in Peck et al., 1989; Ziolkowski & Goldstein, 2008). All five of these studies failed to demonstrate a functional relation because there were less than three demonstrations of an experimental effect. Across all 15 studies, strong effects were seen in both generalization and maintenance. Ten studies (67%) measured at least one form of generalization, and of those studies, all were coded as having strong or mixed outcomes. A different constellation of 10 studies included an assessment of maintenance, and strong or mixed outcomes were reported.

Two findings from this review are particularly relevant to the current study. First, all studies that identified a specific instructional procedure (n=10) as part of the intervention procedures reported strong acquisition outcomes when compared to those that did not report a specific instructional procedure or sequence. The identification of an instructional procedure might be differentially related to strong outcomes; previous research suggests a well-defined, systematically implemented instructional procedure with prompts, error correction, and feedback is more likely to lead to child learning (Sandall & Schwartz, 2002; Wolery, Ault, & Doyle, 1992). Indeed, there is a large body of research in the field of early childhood special education that demonstrates the effectiveness of systematic instructional strategies for children with

disabilities on a variety of target skills (e.g., Alig-Cybriwsky, Wolery, & Gast, 1990; Doyle, Wolery, Gast, & Ault, 1990; Sewell, Collins, Hemmeter, & Schuster, 1998; Wolery, Ault, & Doyle, 1992).

Second, all procedures included a verbal mand. For example, task directions (e.g., "What color is this?") followed by a prompt for the targeted behavior (e.g., "Say blue.") occurred in all 10 studies with a specified instructional procedure. Studies without a specified instructional procedure also indicated mands were part of the intervention. For example, Horn et al. (2000) collected data on a variety of teacher antecedent actions including verbal cues (e.g., the teacher asked the child to count items). Peck et al. (1989) defined teacher prompts as any behavior used to elicit a targeted response from the child. Losardo and Bricker (1994) included planned opportunities for the child to respond (e.g., asked "What do you want?"). Although the lack of specificity reported in this and other studies makes it difficult to discern the exact procedures, it is important to note all studies reported using a mand for a child response.

Commenting Interventions

There is a body of literature on the use of interventions that do not include mands (Leonard, 1981; Girolametto, Pearce, & Weitzman, 1996; MacDonald, 1989; Mahoney & Powell, 1988). In these interventions, the absence of a mand is typically replaced with increased exposures to the target skill by repeated adult verbal productions of the target word. For the purposes of this paper, these interventions will be referred to as *commenting interventions*.

Studies on commenting interventions most often measured the effects of commenting procedures on language production skills, such as phonology, social communication, or grammar. However, a subset of these studies included a measure of vocabulary, which has been

considered a pre-academic skill in research on early literacy development (Howes et al., 2008). Like the instructional approaches described in the studies above, commenting interventions offer an option for embedding instruction into ongoing classroom activities and routines. These studies were not included as part of the Kinder (2009) review because they either used a group experimental design or provided instruction to young children outside of the classroom.

A systematic review of the literature on commenting interventions embedded into ongoing activities and routines in preschool classrooms to target pre-academic skills resulted in the identification of seven studies (Cole & Dale, 1986; Harris & Reichle, 2004; Heal, Hanley, & Layer, 2009; Pemberton & Watkins, 1987; Riches, Tomasello, & Conti-Ramsden, 2005; Wilcox et al., 1991; Yoder et al., 1995). Because only seven studies were found that took place in preschool classrooms, the search was expanded to include commenting interventions that were embedded in ongoing routines and activities in clinics. This search resulted in three additional studies (Ellis Weismer, et al. 1993; Kouri, 2005; Robertson & Ellis Weismer, 1999) for a total of 10 studies. Five of these studies examined the efficacy of commenting interventions and five compared commenting interventions to mand interventions. The five studies that examined the effectiveness of commenting interventions will be described in this section. Those that compared commenting interventions with mand interventions will be described in the following section.

The five studies examining the effects of commenting interventions on children's vocabulary skills used both group experimental and single subject experimental research methodology. Three of the studies used a pretest-posttest randomized control group design (Roberston & Ellis Weismer, 1999; Pemberton & Watkins, 1987; Wilcox et al., 1991), one used a 2 (number) x 2 (spacing) group experimental design (Riches, et al., 2005), and one used a single subject multiple baseline across behaviors design (Harris & Reichle, 2004). Across the 5

studies, 112 children were participants. The majority of the participants were described as late talkers (37%) or typically developing (39%). The studies also included children described as having identified disabilities including specific language impairments (21%), Down syndrome (2%), and moderate cognitive and developmental delays (1%). Participant ages ranged from 20 to 66 months.

The studies were conducted in a variety of settings with different types of implementers. One study, Robertson and Ellis Weismer (1999), trained speech-language pathologists to deliver a commenting intervention during play-based therapy sessions that took place in a clinic. In two studies (Harris & Reichle, 2004; Riches et al., 2005), researchers working in early childhood special education classrooms implemented play sessions. In these studies, it was not clear where the intervention occurred (i.e., inside or outside the classroom). One study (Wilcox et al., 1991) compared the efficacy of commenting interventions delivered in a classroom setting by a teacher to individual intervention delivered by a speech-language pathologist in a clinic playroom. Pemberton and Watkins (1987) delivered a commenting intervention in a small room outside a Head Start classroom, and was the only study to use book reading as the context for instruction.

In all five studies, the target pre-academic skill was vocabulary. However, how vocabulary was measured varied across studies, and was reported either as a broad measure of vocabulary knowledge or a specific number of targeted words learned. Robertson et al. (1999) used parent reported scores on the MacArthur Communicative Development Inventories (CDI; Fenson et al., 1993). Acquisition of targeted words was the only pre-academic dependent measure collected in the other four studies. Variations were reported in the type of words selected as targets and the presentation of the target. In these four studies, 10 to 17 nouns, adjectives, or verbs were selected as targets for each participant. Riches et al. (2005) was the

only study that taught nonsense verbs. Across all five studies, targets were presented verbally, although Harris and Reichle (2004) also taught nouns with a graphic icon.

To be included in this review, the independent variable was a commenting intervention (i.e., delivered repeated exposures to target without a mand). Four of the five studies described the implementation of naturalistic language strategies (described below) as a component of the independent variable. Roberston and Ellis Weismer (1999) used an interactive, child-centered intervention that provided general stimulation emphasizing vocabulary development. Children participated in hands on, theme based, small group activities (i.e., no more than four children) twice a week for 75 minutes. Intervention was incorporated around familiar routines or "scripts" which, when deliberately violated, provided opportunities for the child to use language. However, no mands for language production were given. Instead, techniques such as parallel talk (i.e., verbal description of child's actions), expansions (i.e., repetition of child utterance with addition of one relevant semantic or grammatical information), focused repetitions (i.e., multiple models of the same target word); or recasts (i.e., repetition of child utterance with modification) were used and increased emphasis was placed on targeted words. Feedback was given and related to the communicative utterance (e.g., child who said the word "cookie" was given a cookie).

In another study, Pemberton and Watkins (1987) used one of two specific language strategies (i.e., modeling or recasting) to teach 17 targeted nouns, adjectives, and verbs. Children were randomly assigned to listen to one of two storybooks adapted to include opportunities for either recasts or modeling in the text. No information was provided on whether or not feedback was given when the child produced targets spontaneously. Riches et al. (2005) examined the combination of two training regimes delivered during play sessions on nonsense verb learning.

Interventions varied by the frequency of the verb presentations during play sessions (i.e., 12 or 18 exposures) and the spacing of the sessions (i.e., massed with all exposures delivered on one day or spaced with exposures spread out across four days). During play sessions, the researcher introduced a series of toys and used nonsense verbs to describe the actions while engaging in play with the child. Researcher feedback was not described in the study.

Wilcox et al. (1991) explored effects of variations in the delivery of a commenting intervention. Interactive modeling procedures were either delivered in a classroom setting during regular classroom activities for 3 hours each day by a classroom teacher or during 24 individual play sessions with a speech-language pathologist for 45 minutes. An average of 15 models was given per day in each setting. Interactive modeling procedures began with establishing joint attention by responding to child initiations, following the child's lead, or attempting to interest the child in an activity. Then, the adult labeled the target item that the child focused on (e.g., "cow" if the child is holding a cow). If the child imitated the label, the adult provided an expansion (e.g., "Cow eats."). If the child did not imitate the label, the adult provided the correct label (e.g., child said "doggie", adult said "No, that's a cow.") and continued play without further addressing the child's error. No description of the use of corrective feedback or praise was reported as part of the intervention.

The final study used a commenting intervention that was referred to as aided language stimulation. Harris and Reichle (2004) taught children with developmental delays who were considered functionally nonspeaking (i.e., less than 30 words) to identify spoken and graphic presentations of nouns. A multiple-baseline design across three sets of target stimuli, grouped according to theme (e.g., body parts, food, construction) was used. During play sessions designed to incorporate the themes (e.g., dolls, kitchen, tool bench), the experimenter pointed to

the object, sequentially pointed to the graphic symbol, and then said the target word. This occurred repeatedly during scripted routines with the toy sets. The experimenter delivered non-contingent praise for participation in the session.

Three of the five studies did not include a measure of procedural fidelity for delivery of the commenting intervention (Pemberton & Watkins, 1987; Riches et al., 2005; Wilcox et al., 1991). The remaining two studies measured procedural fidelity using a checklist of treatment components, but the adequacy of the methods of collection varied. Harris and Reichle (2004) collected fidelity on the steps of the procedure during 34-36% of all treatment sessions across all three participants, which is considered acceptable (Horner et al., 2005). Robertson and Ellis Weismer (1999) used a checklist to evaluate how many intervention performance standards occurred during sessions but only collected these data for 5% (n=13) of the play sessions, and did not indicate that data were collected across clinicians. When procedural fidelity was measured, adherence to procedures was high.

Across all studies, participants made gains in vocabulary comprehension and production after receiving commenting interventions. Two ways to measure vocabulary growth were used in the studies: parent reported measures of vocabulary (Roberston & Ellis Weismer, 1999) and targeted words (Harris & Reichle, 2004; Pemberton & Watkins, 1987; Riches et al., 2005; Wilcox et al., 1991). Parents in the Robertson and Ellis Weismer (1999) study completed the CDI as a pre and post measure of their child's global vocabulary skills. The children in the experimental group who received naturalistic strategies during sessions demonstrated significantly greater gains in CDI scores than the delayed treatment control group.

When vocabulary gains were measured by mastery of targeted words, increases also were reported. Pemberton and Watkins (1987) compared two commenting interventions that used

either modeling or recast strategies to teach targeted words during repeated storybook reading. Both methods resulted in significant gains. Riches et al. (2005) reported targeted nonsense word production and comprehension gains for both participants with SLI and those who had normal language. Increased frequency of target words and exposures across intervention sessions were found to have a significant effect for children with SLI when compared with the normal language group. Wilcox et al. (1991) compared commenting interventions delivered in the classroom and in the clinic. Participants increased their use of target words in both settings. All three participants in the Harris and Reichle (2004) study demonstrated gains in the comprehension and production of graphic symbols.

Maintenance was measured in two of the studies. Harris and Reichle (2004) probed comprehension and production of previously acquired target object symbols while providing instruction to subsequent sets of behaviors. Maintenance data were stable, and all three participants maintained high percentages of correct responses. Probes occurred between 11 and 91 days past criterion, providing evidence of sustained effects. A much more limited measure of maintenance was included in the Riches et al. (2005) study. A retention test was given 1 week after the posttest. Although both groups of children showed a loss, children with SLI had poorer retention than those with normal language.

Generalization across settings was formally measured in only one study (Wilcox et al., 1991). Language samples of mother-child interactions were collected in participants' homes mid-way through intervention and after intervention ended. Children in the classroom treatment demonstrated more productive use of target words in their home compared to children receiving individual treatment.

Studies Comparing Commenting Interventions to Mand Interventions

To examine the relative effects of commenting interventions and mand interventions, five studies comparing these two interventions were reviewed (Cole & Dale, 1986; Ellis Weismer et al., 1993; Heal et al., 2009; Kouri, 1995; Yoder et al., 1995). Of these studies, three used group experimental methodology and two used single subject experimental methodology. All three group experimental design studies used pretest-posttest comparison designs, two employed random assignment of participants (Cole & Dale, 1986; Kouri, 1995) and one used a matched groups design (Yoder et al., 1995). Single subject experimental designs included an alternating treatment design (Ellis Weismer et al., 1993) and a multi-element design (Heal et al., 2009). A total of 118 participants were included in these studies. The majority of the participants either had language delays (47%) or overall developmental delays ranging from mild to severe (46%). The remaining participants were either typically developing (5%) or exhibited characteristics indicative of persuasive developmental disorder (2%). Participant ages ranged from 19 to 84 months.

Of the five comparison studies, two delivered commenting interventions in preschool classrooms (Yoder et al., 1995; Cole & Dale, 1986), one delivered the intervention in a small room near the classroom (Heal et al., 2009), and two studies used a clinic play room (Kouri, 2005; Ellis Weismer et al., 1993).The clinic play room was set up similar to a preschool classroom, with toys and materials available for free access. A typical preschool routine was followed.

Measures of pre-academic skills used in comparison studies fell into two categories: global measures of vocabulary and percent correct responses on targeted words during probes. Cole and Dale used the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997) and

Yoder et al. (1995) used both the PPVT and the Expressive One Word Picture Vocabulary Test (EOWPVT; Gardner, 1979). The three studies that measured targeted words varied in terms of the number and type of words that were taught. Ellis Weismer et al. (1993) selected 16 nouns and verbs as target skills. Heal et al. (2009) selected 48 colors and animal names in Spanish. Kouri (2005) measured acquisition of eight nouns.

In the comparison studies, two types of interventions were used: commenting interventions and mand interventions. The commenting interventions were described as follows: modeling (Ellis Weismer et al., 1993), interactive modeling with auditory bombardment (Kouri, 2005), responsive interaction (Yoder, et al. 1995), interactive language instruction (Cole & Dale, 1986), and child-initiated learning opportunities (Heal et al., 2009). Modeling during play and routine activities to provide multiple exposures to target words was used in all the commenting interventions. Other naturalistic strategies used during commenting interventions included parallel talk, self talk, and expansions (Cole & Dale, 1986; Yoder et al., 1995). The number of models used was only reported in two studies (Kouri, 2005; Ellis Weismer et al., 1993) and ranged from 8 to 75. Kouri had a very high number of models because auditory bombardment was used before and after play sessions. During auditory bombardment, the child wore headphones and listened to a recording of repeated productions of the target word while the researcher displayed pictures of each object. Across all comparison studies reviewed, during commenting interventions, children were given opportunities to imitate models through natural interactions with teachers, but were not prompted to imitate models.

Even though children were not prompted to imitate models when a commenting intervention was used, spontaneous production of targets occurred, providing an opportunity for the adult to deliver feedback. No studies provided corrective feedback when a spontaneous

production was incorrect. When a spontaneous production was correct, the type of feedback received varied across studies. Heal et al. (2009) provided verbal praise for correct responses, (e.g., "That's right! That is azul!"). Teachers in the Cole and Dale (1986) study avoided using verbal praise, and instead used natural consequences when a spontaneous target was correct (e.g., giving a cup when child asks for a cup). Two studies (Ellis Weismer et al., 1993; Kouri, 2005) delivered a semantically related expansion when the child produced a target word (e.g., child says "doll", adult says, "goodnight doll"). Ellis Weismer et al. acknowledged the target was correct before delivering the expansion (e.g., saying "uh huh" or "yes" before delivering expansion). Yoder et al. (1995) provided contingent imitation of communicative attempts including target productions (e.g., child says "doll", adult picks up doll and says "doll"). Mand interventions were described as follows: modeling plus evoked productions (Ellis Weismer et al., 1993), mand-elicited imitation (Kouri, 2005), milieu teaching (Yoder et al., 1995), and DISTAR direct instruction (Cole & Dale, 1986). Two mand interventions were used in Heal et al. (2009). Both included mands following child-initiated opportunities, and one also included mands following teacher-initiated opportunities. Three studies preserved the context and modeling used in the commenting intervention, but embedded mands into the ongoing natural interactions (Heal et al., 2009; Ellis Weismer et al., 1993; Yoder et al., 1995). For example, in the Yoder et al. study, teachers used models to prompt responses within the context of responsive interaction. In the context of play, teachers would a deliver a mand (e.g., "What do you want?") followed by a prompting procedure to elicit production of the targeted word. Two interventions varied by more than just the inclusion of the mand (Cole & Dale, 1986; Kouri, 2005). Kouri (2005) eliminated the auditory bombardment component that was delivered during the commenting intervention Cole and Dale used materials from the *DISTAR Language I Program* (Englemann & Osborn,

1976) during small group instructional sessions focused on teaching pre-determined language goals. Activities were teacher-initiated and relied heavily on children responding to teacher prompts to produce appropriate responses.

Feedback to children following correct and incorrect responses to mands varied across the mand interventions used in the five comparison studies. Both mand interventions used in the Heal et al. (2009) study included praise following correct responses and corrective feedback following incorrect responses. Cole and Dale (1986) also delivered verbal praise for correct responses, but did not report whether feedback was given for incorrect responses. Ellis Weismer et al. (1993) and Yoder et al. (1995) both delivered feedback on the correctness of the response. Ellis Weismer et al. also included a model of the target in the feedback (e.g., "Right, it's the pen."). Yoder et al. added an expansion or extension of the child's utterance in the feedback. Functional consequences were used in two studies (Kouri, 2005; Yoder et al., 1995). In both studies, an activity was withheld until a target was produced.

Three of the five comparison studies included a measure of procedural fidelity. Ellis Weismer et al. (1993) analyzed videotape recordings of 14% of the sessions, with half of the fidelity sessions randomly chosen from each instructional condition. Three components of the intervention were examined (i.e., number of models of the target words, number of evoked productions, frequency of feedback). Adherence to the procedures was high. Yoder et al. (1995) collected procedural fidelity in two ways. First, research staff completed treatment feedback sheets. These were given to the teacher as a method of feedback and included checklists of critical components of the intervention that were observed in practice. In addition, 20 times during the study, videotaped samples were coded in detail. These were also used to deliver feedback to the teachers regarding their adherence to intervention procedures. Average correct

implementation of responsive interaction procedures was 70% (SD 16.9), and for milieu teaching, it was 69% (SD 19.8). The high standard deviations indicate that across settings and observations, teachers varied in their ability to deliver the instruction with high fidelity.

Overall, commenting interventions and mand interventions showed comparable results. Cole and Dale (1986) and Yoder et al. (1995) did not show statistically significant differences in vocabulary growth between the two interventions. However, Yoder et al. reported differential effects when the developmental level of the child was taken into consideration. Children with lower language scores at pretest had the greatest vocabulary gains in milieu teaching, while the children with higher receptive age equivalents benefited more from responsive interaction. Child characteristics may have also influenced results in the Ellis Weismer et al. (1993) study. They reported the commenting intervention was more effective than the mand intervention for one participant, the mand intervention was more effective than the commenting intervention for another participant, and neither produced gains for a third participant.

Advantages for mand interventions over commenting interventions were seen in two studies (Heal et al., 2009; Kouri, 2005). Kouri found the number of target words that met acquisition criteria were significantly higher in the mand intervention than the commenting intervention. Several other measures of performance (i.e., number of spontaneous words produced, total treatment productions, and number of sessions before target words were used spontaneously) were not significantly different across treatments. Although some comparable gains were made, the commenting intervention did not produce the same frequency or consistency of target word use as the mand intervention. Heal et al. found the strategy that included commenting plus mands after both child-initiated and teacher-initiated opportunities was the most efficacious in terms of the highest number of correct responses, fewest number of

sessions to reach criterion, and highest post-test scores. However, they noted children who received the commenting only strategy made some gains in target word production, and had higher post test scores than children who received the commenting plus mands after child-initiated opportunities.

Only one comparison study measured generalization. Kouri (2005) measured generalization of target words by collecting home language samples two weeks after the final treatment. Participants who received the commenting intervention (i.e., modeling plus auditory bombardment) produced more target words in the generalization setting than those in the mand-elicited intervention. Also, participants who received the commenting intervention spontaneously produced words in the home that were not produced in the clinic. This finding was not reported for participants who received the mand intervention. None of the comparison studies measured maintenance.

Summary and Limitations

The Kinder (2009) review found support for the effectiveness of embedded instructional approaches including a mand, for teaching pre-academic skills. The literature on commenting interventions (i.e., those that do not include a mand and include multiple exposures to the target skill) demonstrates that these interventions can be effective for teaching vocabulary. No commenting studies were found that targeted pre-academic skills beyond teaching vocabulary. When commenting interventions were compared to mand interventions, little to no difference was seen in acquisition outcomes. When differences were found, child characteristics appeared to be a factor.

Data on the generalization and maintenance of target skills were included in the majority (67%) of the studies included in the Kinder (2009) review. When compared to studies on mand interventions, fewer commenting studies measured generalization (20%; *n*=1), and maintenance (60%; *n*=3) of vocabulary gains. In studies that compared commenting interventions to mand interventions, only one (20%) measured generalization and none measured maintenance. Overall, when measured, studies that used a commenting intervention and those that used a mand intervention demonstrated positive generalization and maintenance outcomes. The one study that compared generalization across both interventions reported stronger generalization across setting outcomes for the commenting intervention as compared to the mand interventions.

Several limitations are noted across the three reviews. First, only one pre-academic skill (i.e., vocabulary) was taught in studies that included a commenting intervention. Thus, the effects of commenting interventions on other pre-academic skills are unknown. Second, few commenting studies measured generalization of acquired skills. Limited data suggest commenting interventions may be more effective at promoting generalization when compared to interventions with a mand. However, this research is limited by the number of studies. Third, the comparison studies did not report maintenance data, and thus little is known about the relative effects of the two types of interventions on the maintenance of outcomes. Fourth, although the delivery of the commenting interventions was typically described with replicable precision (e.g., model, expansion, recast, auditory bombardment), few studies reported the number of comments delivered or assessed procedural fidelity. These limitations warrant further research on the use of interventions that use mands, comments, or some combination in the context of play with children and the effects on acquisition, generalization and maintenance.

Purpose of the Current Study

The current study expands the literature on embedded instruction in several ways. First, the current study examined the effects of a commenting intervention on an academic skill other than vocabulary (i.e., letter naming). Letter naming had not been taught in previous studies of commenting interventions. Second, the current study compared a commenting intervention to an intervention that combined comments and mands, by adding the use of constant time delay (CTD; Wolery, Ault, & Doyle, 1992), an instructional procedure that includes a request for the child to respond, (i.e., "What letter is this?"). Third, the current study included multiple measures of efficiency. Fourth, the current study included more extensive measures of maintenance and generalization than had been used in previous comparison studies. Fifth, the study examined a systematic application of commenting including careful measurement of procedural fidelity.

The purpose of the current study was to evaluate the relative effects of two interventions to teach a pre-academic skill. In both interventions, instruction was delivered during play activities designed to maintain child participation, interest, and engagement. Descriptive comments were used in both interventions to expose children repeatedly to targeted letter names in the context of salient information about that letter. In one intervention, an instructional procedure (i.e., CTD) was implemented in addition to the commenting. These trials always included a mand. The following research questions were addressed in this study:

1. Does an intervention that includes a systematic application of descriptive comments result in acquisition of expressive letter naming?
2. Does an intervention that includes a systematic application of descriptive comments plus constant time delay result in acquisition of expressive letter naming?

3. Does an intervention using descriptive comments or an intervention using descriptive comments plus constant time delay result in more efficient acquisition of expressive letter naming?

4. Does an intervention using descriptive comments or an intervention using descriptive comments plus constant time delay result in greater generalization and maintenance of expressive letter naming?

CHAPTER II

METHODS

Child Participants

Four preschool-age children were recruited from two early childhood centers in a southeastern city. Children were included based on the following criteria: (a) a chronological age from 3 to 5 years at the beginning of the study, (b) history of consistent attendance in school (i.e., 80% attendance for previous month), (c) inability to name expressively at least 10 letters of the alphabet when the study began, and (d) ability to maintain attention in a one-on-one activity with an adult with minimal supports (e.g., visual presentation of the task, no more than three reminders or redirections) for at least 5 minutes. Inclusion criteria related to age and attendance records were confirmed by each child's classroom teacher. Screening procedures were used to determine if the child met the inclusion criteria related to attention and letter knowledge. Screening data were also used to select target stimuli.

All screening procedures were conducted by the PI or a graduate assistant familiar with the procedures. Screening data are shown in Table 1. The first screening task was a rapid letter naming (RLN) task. The child was presented with four rows of capital letters in random order and was instructed to name all the letters he could in 1minute. The PI revealed one row of letters at a time and pointed to each letter. This was done to determine if there were 10 letters the child did not name expressively. Based on the results of this task, a list of eligible target letters was created. Eligible target letters included letters the child did not name expressively in the RLN task and were not in the child's first name or last initial. Next, expressive knowledge, receptive

Table 1

Child Screening Results

Child	RLN score	Eligible letters	Named expressively [a]	Identified receptively [b]	Matched [c]
Jamal	4	18	0	2	18
Reggie	8	15	5	6	15
Keandre	3	20	3	5	20
Mikey	3	20	0	0	14

Note. RLN= rapid letter naming.
[a]Named correct letter 1/2 or 2/2 trials. [b]Identified correct letter 3/4 or 4/4 trials. [c]Matched letter correctly 2/2 trials.

knowledge, and the ability to match the eligible letters were screened with a series of tasks. Expressive knowledge was tested again to confirm RLN task results. The PI showed cards with written letters to the child, and asked the child to either name the letter (two trials presented), point to the letter (four trials presented), or match it to a letter that looked exactly the same (two trials presented). A list of the letters the child matched but did not name was created. If this list had less than 10 letters, the child was excluded from the study.

To provide a description of the developmental status of participating children, three measures were administered. The *Mullen Scales of Early Learning* (Mullen, 1995) was used as a measure of early intellectual development and school readiness. The measure provides age equivalent scores in each domain (i.e., visual reception, fine motor, receptive language, expressive language) as well as an overall composite score, which represents an average of the child's performance on the Visual Reception, Fine Motor, Receptive Language and Expressive Language scales. This instrument was standardized on a nationally representative sample across gender, race/ethnicity, community size, and socioeconomic status. The sample included typically developing, English speaking children from all regions of the United States. The *Preschool Language Scale, Fourth edition* (PLS-4; Zimmerman, Steiner, & Pond, 2002) measures children's expressive and receptive language skills. Norms used in this measure were based on a large, diverse sample of children. Based on expert review, test items were reported to be appropriate for children from different social economic status (SES) and ethnic groups. The PLS-4 provides an auditory comprehension score and an expressive communication score. The *Test of Early Reading Ability, Third edition* (TERA-3; Reid, Hresko, & Hammil, 2001) provides a direct measure of early reading skills. The normative sample was selected to be representative of the United States population, and studies on the measure showed an absence of racial and ethnic

bias. The TERA-3 provides a measure of overall reading ability (i.e., Reading Quotient) which is derived from standardized scores on three subtests: Alphabet, Conventions, and Meaning. The alphabet subtest measures the child's knowledge of the alphabet and sound-letter correspondence. The conventions subtest measures the child's familiarity with the conventions of print. The meaning subtest measures the child's ability to comprehend the meaning of printed material.

Information regarding each child's approach to engagement in classroom activities, including those targeting literacy skills, was collected via a written survey (Appendix A) completed by each child's teacher. The survey included questions about the child's preferred classroom activities, level of engagement in instructional activities, and typical interaction with literacy materials. All participants were males and had no identified disabilities. Jamal, Keandre, and Reggie were African-American and Mikey was Caucasian. Specific information about each child is described below. Testing information is shown in Table 2.

Jamal. Jamal was 5-years-3-months old when the study began. He identified four letters in the RLN task. He scored 3 to 21 months below age level on the Mullen subtests, resulting in an early learning composite score of 72, which falls in the below average range. He scored highest in the expressive language subtest and lowest in the receptive language subtest. On the PLS-4, he scored 6-10 months below age level, with a higher auditory comprehension score than expressive communication. On the TERA-3, Jamal received an overall reading quotient of 79, which was considered poor. He was 1 year below age level for both conventions and meaning and 2 years below age level for alphabet knowledge. His teacher, Cornelia, reported his favorite activity in the classroom was building at the block center. She also reported he often did not engage in group or individual book reading and large or small group literacy activities.

Table 2

Testing Information

		Mullen[a]					PLS-4[b]		TERA-3[c]			
Child	Age	Early Learning Composite	Visual Recept.	Fine Motor	Recept. Lang.	Exp. Lang.	Aud.	Expres.	Reading Quotient	Alpha.	Conven.	Meaning
Jamal	5-3	72 (Below Average)	4-6	4-4	3-6	5-0	4-9	4-5	79 (Poor)	3-3	4-3	4-3
Reggie	5-4	68 (Very Low)	4-0	4-7	3-8	4-5	5-0	5-9	74 (Poor)	3-3	4-3	4-9
Keandre	4-8	80 (Below Average)	4-3	3-8	3-2	3-4	4-3	4-3	85 (Below Average)	4-3	3-3	3-6
Mikey	4-0	92 (Average)	4-2	3-4	3-5	4-0	3-9	4-3	72 (Poor)	3-3	3-3	3-3

Note. Unless otherwise noted, all scores are given as age equivalent (year-month).
[a]*Mullen Scales of Early Learning* (Mullen, 1995).[b]*Preschool Language Scale, Fourth edition* (Zimmerman, Steiner, & Pond, 2002).[c]*Test of Early Reading Ability, Third edition* (Reid, Hresko, & Hammil, 2001).

According to Cornelia, Jamal typically refused to participate in activities that were challenging and would cry or not respond to attempts to engage him.

Reggie. Reggie was 5-years-4-months old when the study began. He named eight letters in the RLN task. He scored 9 to 20 months below age level on the Mullen subtests, resulting in an early learning composite score of 68, which falls in the very low category. He scored highest on the fine motor subtest and lowest on the receptive language subtest. On the PLS-4, he scored 4 months below age level on the auditory comprehension subtest, but 5 months above age level on the expressive communication subtest. On the TERA-3, Reggie received an overall reading quotient of 74, which was considered poor. He was 25 months below age level for alphabet knowledge, 13 months below age level for conventions, and 7 months below age level for meaning. His teacher, Cornelia, reported his favorite activities in the classroom were block center activities such as cars and trucks. She also reported he had more trouble staying engaged in large group literacy activities than during individual book reading or small group literacy activities. According to Cornelia, Reggie typically tried to participate in literacy activities but was hesitant when asked to respond and seemed unsure about his abilities. She felt he would benefit from more one-on-one activities with an adult.

Keandre. Keandre was 4-years-8-months old when the study began. He named three letters in the RLN task. He scored 5 to 18 months below age level on the Mullen subtests, resulting in an early learning composite score of 80, which is in the below average range. He scored highest in the visual reception subtest and lowest in the receptive language subtest. On the PLS-4, he scored 5 months below age level on both the auditory comprehension and expressive communication subtests. On the TERA-3, Keandre received an overall reading quotient of 85, which was considered below average. He was 5 months below age level for alphabet knowledge

and more than a year below age level for conventions and meaning. His teacher, Cornelia, reported his favorite activity in the classroom was playing at the computer. She also reported he usually stayed engaged when reading books on his own or while in small group literacy activities, but he had trouble staying engaged in large group book reading and literacy activities. She was concerned about his ability to comprehend information, and indicated he would often talk about events that were not related to the story being read. Cornelia also expressed concerns regarding his inability to attend to instructional tasks.

Mikey. Mikey was 4-years old when the study began. He named three letters in the RLN task. He scored at or slightly above age level in the expressive language and visual reception subtests of the Mullen. His fine motor and receptive language scores fell about 6 months below age level. His early learning composite was 92, considered average. On the PLS-4, he scored 3 months below age level on the auditory comprehension subtest and 3 months above age level on the expressive communication subtest. On the TERA-3, he received an overall reading quotient of 72, which is considered poor. He was 9 months below age level for all three subtests. His teacher, Jocelyn, reported his favorite classroom activities were dressing up and playing with cars. She reported he usually stayed engaged when participating in play or one-on-one literacy activities with an adult, but did not remain engaged when reading or looking at books on his own. According to Jocelyn, Mikey also did not stay focused and engaged during large group activities.

Settings

Three of the child participants (i.e., Jamal, Keandre, &Reggie) attended a pre-kindergarten classroom located in a non-profit urban child care center serving low-income

families. The center was located near federally funded housing projects. There were 10 classrooms for children ages 2 - 5 years old. Class sizes varied, but typically averaged about 18 children. School programming emphasized literacy skills, and a literacy coach provided support to the teachers in the form of testing children, providing materials, and providing staff development. Cornelia was the lead teacher in the classroom these three children attended. On the classroom observation items of the *Early Language and Literacy Classroom Observation (*ELLCO; Smith & Dickinson, 2002), Cornelia's classroom received an average score of 4.57, indicating near exemplary literacy practices. Cornelia reported she used the *DLM Early Childhood Express Curriculum* to plan all classroom activities. The activities Cornelia reported using most often to help children learn letters were: discussing vocabulary words daily, pointing out print in books, and planning early word recognition activities. In terms of literacy activities, she estimated children in her classroom spent the most time reading books on their own (more than 45 min) and the least amount of time participating in small group literacy activities (less than 15 min). Cornelia reported she believes children learn letters best by having print and writing materials always available to them, watching adults write, having other environmental print in the classroom, and doing small group journal work related to the current theme.

Mikey attended an inclusive child care center located on the campus of a major university. Eight classrooms provided year round programming for infants through 5-year-olds. Class size varied, but most classrooms had about 12 children, including 3 to 4 children who had an identified disability. When Mikey began the study, Jocelyn was his teacher in a classroom for 3-year-old children. During the study, he transitioned into a new classroom. This was part of the program's cycle of promoting students to the next age level near the end of the summer. During the comparison phase, Julia became Mikey's teacher in a classroom for 4-year-old children. This

occurred after a 3-week break. Because she was Mikey's teacher at the beginning of the study, Jocelyn completed all his study procedures (i.e., demographic survey, generalization across adults sessions). However, the ELLCO was completed in both classrooms, and results were comparable. On the Classroom Observation items, Jocelyn's classroom received an average score of 2.86 and Julia's classroom received an average score of 3.14. A score of 3 classifies their language and literacy environments as basic.

All study procedures were conducted in a one-on-one arrangement with the child and an adult. In both child care centers, all screening, probe, intervention, and generalization across materials sessions took place outside the classroom. Generalization across adults sessions took place in the classroom.

For Jamal, Reggie, and Keandre, screening and morning intervention sessions typically occurred at one of several adult or child-sized tables and chairs located in the hallway adjacent to the classroom. Volunteers frequently used these areas to interact with individual children around academic tasks or play activities. Afternoon intervention sessions typically occurred in a large conference room with an adult-sized table and chairs. Probe sessions typically occurred at an adult-sized table and chairs in the teacher's lounge. The conference room and teacher's lounge were used whenever available to minimize distractions from hallway traffic or other potential interferences. Generalization across materials sessions took place in locations that offered comfortable places for reading including: a designated library area in the corner of a multi-purpose room; in the classroom on cushioned, child-sized chairs or plush carpeting; or on an adult-sized couch in the teacher's lounge. Generalization across adults sessions took place at a child-sized table in the classroom.

For Mikey, screening, probe sessions, and morning intervention sessions typically occurred at a child sized table and chairs in a small play room used by therapists or for after school care. Afternoon intervention sessions typically occurred in a classroom that was not in use at the time of the session. Generalization across materials sessions took place in locations that offered comfortable places for reading including an empty classroom on cushioned, child-sized chairs or sitting next to each other on plush carpeting. Generalization across adults sessions took place sitting across from each other on a carpeted area in the classroom. On the few occasions these locations were not available, alternate locations included a picnic bench on the playground, a small area in the corner of the gym, or an unused corner of the classroom (e.g., the dramatic play area during snack time).

Materials

Six types of materials were used in this study: letter cards, reinforcers, letter activities, activity choice boards, picture books, and data collection materials. All stimulus materials featured capital letters to reduce potential errors related to the orientation of the letter rather than the child's knowledge of the letter. For example, the lower case letters b, p, d, and q are the exact same combination of shapes oriented differently, but the capital forms of these letters do not have similar characteristics.

Letter cards were used in screening and probe sessions. Letters were hand written with a thick black marker on 10.2 x 15.2 cm white index cards. Finger puppets and a glitter filled wand were used as pointers during probe sessions to increase engagement. Child sized sunglasses without lenses (referred to as "letter finding glasses") were used to increase engagement in

screening and probe sessions. Stickers were given as reinforcers for participation in screening and probe sessions.

Six different activities that featured uppercase letters as a salient component were used during the comparison phase of the study. Activities included the following commercially produced or researcher created games: (a) *Don't Break the Ice!* by Milton Bradley, (b) *Penguin Pile Up* by Ravensburger Games, (c) *Super Spiral Speedway* by Fisher Price, (d) fishing for letters, (e) cooking with letters, and (f) buried letter treasure. These six letter activities were used in intervention sessions across both interventions on a systematically rotating schedule. An activity choice board was created with laminated photographs of key materials used in each activity mounted with Velcro to a laminated piece of colored cardstock. Materials in all activities were adapted to include the child's target letters. Each child's target letters were written with black marker on 1.9 cm white round labels and affixed to interchangeable pieces. Target letters were written directly on the *Don't Break the Ice* game pieces because they already had a flat white background suitable for writing. Letters were comparable in size and shape across all intervention and generalization across adults activities.

As the child became less interested in an activity, new materials were added or variations on the activity were introduced. However, the activity and the pieces containing the target stimuli remained the same throughout. For example, when playing cooking with letters, colored plastic dominos with target letter labels were always the objects that were cooked. However, a variety of supplemental materials (e.g., popcorn bags and pretend grated cheese, cookie sheet and oven mitt, french fry basket and toy salt shaker) was introduced as children required more prompts to participate in the task.

During Keandre's last nine sessions, three additional activities were added to the rotation of materials in an attempt to support his engagement in the activities: (1) *Topsy Turtle* by Hasbro, (2) *Football Knock Down* by SportsTime, and (3) *Tri Level Garage* by Matchbox.

During generalization sessions, a letter activity and picture books were used. During generalization across adults sessions, Disney's Toy Story *Alien Grab* game was used. During generalization across materials sessions, five picture books were used. These books featured capital letters with one or more distinguishing features such as a larger font size or color interspersed throughout the text. Adaptations were made to the books as needed to add distinguishing features to selected capital letters (i.e., writing large capital letters on the pages with colored markers), eliminate potentially confusing stimuli (i.e., marking through a lower case letter printed next to the upper case letter), and provide multiple opportunities for letter naming (e.g., drawing a bowl of alphabet soup on the back page of a book about food, adding paper flaps with letters to existing illustrations).

Data collection materials included data sheets and procedural fidelity checklists. During probe, intervention, and generalization sessions, a stop watch or kitchen timer was used to recorded session length. In addition, a 15.24 cm, free-standing, double-sided stop and go sign was used to indicate to the child the beginning and ending of each session. A digital audio recorder was also used during intervention sessions to confirm data collected live were accurate.

Experimental Design

An adapted alternating treatments design (Sindelar, Rosenburg, & Wilson, 1985) replicated across four participants was used to compare the effects of two intervention conditions, descriptive comments (DC) versus descriptive comments plus trials (DC+T), on

expressive letter naming. This design allows the comparison of two instructional treatments by rapidly alternating instruction on two equivalent sets of target stimuli while probing for learning on a third (control) set of targets. The treatment that results in acquisition the quickest is identified as the superior treatment. Experimental control is demonstrated by the absence of learning in each participant's control set. Applying the superior treatment to the control set after the comparison phase serves as an intra-subject replication (Wolery, Gast, & Hammond, 2010).

General Procedures

Prior to baseline, consented children were screened using a rapid letter naming test to ensure they met inclusion criteria. Next, expressive, receptive, and matching tasks were conducted to confirm other inclusion criteria and select stimuli for the comparison phase. In addition, three developmental assessments were conducted with each child. Finally, the Early Language and Literacy Classroom Observation (ELLCO; Smith & Dickinson, 2002) was conducted in each classroom as a descriptive measure of the classroom environment related to language and literacy teaching.

During baseline, at least three probe sessions were conducted and visual analysis was used to determine stability of child responses. Three generalization across materials and two generalization across adults sessions also were conducted during baseline. During the comparison phase, probe sessions were conducted each day prior to the first intervention session. Barring absence or interference from unforeseen events, two intervention sessions occurred every school day, one in the morning and one in the afternoon. Interventions (DC or DC+T) were counterbalanced for time of day and order of adjacent sessions. Intervention sessions continued until the predetermined learning criterion was met for one of the sets. Criterion was three

consecutive probe sessions with 100% unprompted correct responses. After criterion was reached in one set, the intervention sessions for that set occurred every fourth session, regardless of time of day. When criterion was reached in the other set, a maintenance condition was implemented. Intervention for the second set continued for 1.5 times the number of sessions required to reach criterion for the first set. If criterion was not reached by that point, intervention ceased. During maintenance, intervention sessions were not conducted, but probe sessions and generalization sessions continued. Finally, the superior intervention (i.e., the intervention that required fewer intervention sessions to meet criterion) was used to teach the control set. Once each week throughout the study, a generalization across materials session was conducted. Generalization across adults sessions occurred twice during maintenance. Maintenance probe sessions were conducted for the mastered sets for 1 to 5 weeks following the end of intervention.

Measurement

Probe data. Probe sessions and generalization across materials sessions were conducted in all three conditions: baseline, comparison, and maintenance. Generalization across adults sessions were conducted in baseline and maintenance. Data were collected live during all probe and generalization sessions. The PI collected all primary data using an event recording paper/pencil observational system. Child responses were recorded by the PI by either circling the correct code or making a tally mark on the appropriate data sheet (Appendix B).

During all probe and generalization sessions, the adult secured the child's attention and gave the task direction, "What letter is this?" Possible responses included: unprompted correct, unprompted error, and no response. An unprompted correct was scored when the child said the correct answer within 4 s of the task direction. An unprompted error was scored when the child

responded with an incorrect or unrelated answer within 4 s of the task direction. No response was scored when the child did not say anything within 4 s of the task direction. The first utterance the child made was coded. An exception was made if the first utterance the child made was one of the probe session rules (e.g., "Think, think, think… letter X!") or one of the descriptive comments (e.g., "Criss cross, it's the letter X!"). The dependent measure in the study was the percentage of unprompted correct responses following the task direction, "What letter is this?" during probe sessions. This was calculated by dividing the number of unprompted correct responses by the total number of trials presented in the session and multiplying by 100.

Intervention session data. Data were collected during all intervention sessions in both interventions. Data were collected live and confirmed with audiotape review. During the descriptive comments sessions (DC; described below), each time the child said the target letter it was recorded on the data sheet, and each was coded as spontaneous or imitative. A spontaneous target was defined as any utterance made by the child during an intervention session that included the correct name of the target letter and occurred at least 10 s after the PI said the same letter. A spontaneous target was only recorded if it was clear the child was correctly referring to the letter stimulus. For example, if during the session the child said, "The letter B car is my favorite," and he was not looking at or playing with a car with the letter B, nothing was scored. However, if he made the same utterance while holding the car with a B on it, it was coded. An imitative target was defined as any utterance made by the child during the intervention session that included the target letter name and occurred 9 s or less after the PI said the same letter. An imitative target did not require a correct reference to a target letter stimulus, because the PI's utterance served as the reference. For example, if the PI said, "You have the letter B on your penguin!" and 3 s later the child said "B!", an imitative target was recorded regardless of what

the child was looking at or manipulating the relevant material. The total number of child spontaneous and imitative targets for each letter was calculated for each session.

During the descriptive comments plus trials sessions (DC+T; described below), the following child behaviors were coded: unprompted correct, unprompted error, no response, prompted correct, prompted error, spontaneous targets, and imitative targets. The same definitions for unprompted correct, unprompted error, and no response described above for the probe sessions were used for DC+T sessions. In addition, a prompted correct was scored when the child said the correct answer within 3 s of the model prompt (e.g., "Say B"). A prompted error was scored when the child said the incorrect answer within 3 s of the model prompt. The same definitions for spontaneous and imitative targets were used as described above, and were only coded outside the time delay trials. A trial began as soon as the task direction was delivered and ended after the PI delivered descriptive praise or feedback. If the child imitated the feedback, it was coded as an imitative target.

Interobserver Agreement

A graduate assistant familiar with study procedures collected interobserver agreement data for a minimum of 25% (range 25 - 100%) of each child's probe sessions in each condition and in a minimum of 25% (range 26.32 - 55%) of all intervention sessions for each child. The graduate assistants and the PI established reliability before collecting data for the study. One graduate assistant was trained prior to the beginning of the study using observations in practice settings. Other observers were trained through observations of study sessions, but their data were not used until they were reliable.

Interobserver agreement for child responses during probe and intervention sessions was calculated using a point-by-point formula (Kazdin, 1982). The number of agreements was divided by the number of agreements plus disagreements and multiplied by 100. During intervention sessions, the PI and the graduate assistant independently recorded spontaneous or imitative targets. The percentage of agreement scores for child targets was calculated using a total agreement approach. The total number of spontaneous targets for each observer was totaled; the smaller total was divided by the larger total, then multiplied by 100. The same process was applied to the imitative target data.

Procedures

Identifying target letters and creating sets. Use of the adapted alternating treatment design requires the identification of equally difficult sets of target behaviors. Careful attention was given to the selection of target letters and creation of sets to maintain the integrity of the design. Equal difficulty was addressed by thoroughly screening participants to determine related letter knowledge, establishing rules for inclusion in sets, and matching and randomly determining stimuli for each set.

Using the screening data, a series of steps was implemented to select sets of target stimuli. First, a list of letters the child did not name (0/2 trials), did match (2/2 trials), and clearly articulated was compiled. From this list, receptive identification data were taken into consideration to ensure equivalency across sets. Only letters that were receptively identified in no more than 2/4 trials were included in letter sets. Using the letters that were identified in the previous steps, a list of every possible letter pairing was created. The pairs were examined and

eliminated if letters were next to each other in the alphabet, letters resembled each other in shape (e.g., W and M), or letters had similar sounds (e.g., B and V).

The remaining letter pairs were assigned to sets as follows. Each pair was assigned a number. Three numbers were selected by a random number generator. The pair assigned the first randomly generated number was designated as set 1. Set 1 letters were taught using the DC intervention. The pair assigned the second randomly generated number was designated as set 2. Set 2 letters were taught using the DC+T intervention. The pair assigned the third randomly generated number was designated as set 3, the control set. This process was used for Jamal, Keandre, and Reggie, who began the study at the same time. After the sets were selected for all three participants, all nine sets were evaluated to ensure no letter was assigned to the same set across all three participants. One such case was detected, and the sets were randomly reassigned. The target letters, receptive screening information, and set assignments for each participant are shown in Table 3.

Two issues with the letters assigned to set 3 (control) for Jamal and Reggie occurred during baseline and were addressed as follows. Jamal named the letter Q during the generalization across adults probe in baseline. Because he had not named it during probe sessions in baseline, it remained in the set as an eligible letter. Reggie consistently named one of the control letters (U) during baseline. In the fourth probe session, the letter was removed, and a replacement letter (M) was randomly chosen from his list of eligible letters.

The procedures for selecting Mikey's target letters varied slightly from those used for the other three children. Procedures for screening and identifying letter sets were the same as described above with the following exception. Based on informal information from the first three children, the first letters of the names of the children who were in his class were noted (C, X).

Table 3

Participant Target Letters and Receptive Screening Information

	Set 1 (DC)				Set 2 (DC+T)				Set 3 (control)			
Child	Letter 1	ID recept.	Letter 2	ID recept.	Letter 1	ID recept.	Letter 2	ID recept.	Letter 1	ID recept.	Letter 2	ID recept.
Jamal	D	1/4	F	1/4	G	0/4	K	1/4	Q	1/4	T	0/4
Reggie	J	2/4	V	1/4	F	0/4	W	2/4	H	1/4	M	2/4
Keandre	U	1/4	Y	1/4	F	0/4	J	1/4	Q	0/4	S	1/4
Mikey	P	0/4	X	0/4	D	0/4	Q	0/4	R	1/4	W	1/4

Note. DC = descriptive comments; DC+T = descriptive comments plus trials.

These letters were not excluded, but they were not paired together. Also, pairings containing any of the six letters used in each of Jamal, Keandre, and Reggie's sets were eliminated prior to randomly assigning that set for Mikey. This was done to provide as much variation as possible in the letters that were assigned to each set across the four participants.

Probe procedures. After the sets were created, probe sessions began. At the beginning of each probe session, the children were reminded of the rules. Rules were: (1) look at the letter; (2) think, think, think; and (3) say the letter in a loud voice. During probe sessions, the PI secured the child's attention, presented each target letter printed on an index card, and delivered the task direction, "What letter is this?" Each letter was presented 3 times for a total of 18 trials (3 sets of 2 letters = 6 letters X 3 trials = 18). To provide opportunities for success, three known letters were presented two times each for a total of six trials. Thus, 24 trials were presented in each probe session. No data were collected on known-letter responses. Cards were randomly shuffled and presented. After delivering the task direction, the PI waited up to 4s and then recorded the child's response on a data sheet. Praise was given for correct responses (e.g., "You got it right!" or "Good job!"), and neutral comments were given after errors and no responses (e.g., "Ok."). Encouragement was given throughout the session as needed to maintain children's attention (e.g., "Keep going – only a few more!" or "You're working so hard, you can do it!"). At the end of a probe session, children received a sticker.

On occasion, children required multiple redirections to attend to letter stimuli during the probe sessions. When this occurred, the PI gave the child the option to use a prop. Two types of props were used: glasses and pointers. Various pairs of child-sized sunglasses with the lenses removed were referred to as "letter finding glasses". Another prop was a plastic tube filled with glitter with various finger puppets placed on the end. The child was given the option to choose a

finger puppet to put at the end of the tube and use it to point to the letter on the card as he named it. These supports were used with Jamal, Keandre, and Reggie. Keandre used the props most frequently. Mikey did not require use of the props to maintain attention to the letter cards.

Comparison phase. Once baseline data were stable, the comparison phase began. Probe sessions continued during this phase. Probe sessions were always the first study activity that occurred each day. Procedures for probe sessions during the comparison phase were identical to the procedures used during baseline, except the control set (set 3) was included only intermittently (i.e., every third probe session instead of every session). After criterion was reached on one letter set, known letters were removed from the stimuli to reduce the time spent in probe sessions.

During the comparison phase, probe procedures were adapted for individual children as follows. To ensure Jamal was attending to the discrete features of the letter stimuli, a matching task was added before the task direction ("What letter is this?") was delivered. This procedure was used for probe sessions 16 - 31 and occurred as follows. Two cards with letters from the set were presented face up on the table. Jamal was shown the target stimuli and was told to match. After he placed the target stimuli on the correct card, the PI pointed to the stimuli and delivered the task direction. All other procedures remained the same. Keandre required multiple redirections to attend to the letter stimuli before responding to the task direction. To ensure he was attending to the stimuli before responding, the same matching task procedures used for Jamal were used for Keandre from session 18 to the end of the study. Beginning with probe session 24, only Keandre's set 1 and set 2 stimuli were presented, and the known letters were removed to reduce the number of trials. In addition, the intermittent schedule for the control set was increased from every 3 sessions to every 6 - 8 sessions. A scheduled break was implemented

for Jamal and Keandre beginning at probe session 17 and 18, respectively. After half of the trials had been delivered, the child was allowed to play with a small toy for a few minutes. During probe sessions when a break was used, the PI drew a visual schedule of the sequence of the probe session activities (i.e., letter cards, break, letter cards, sticker) and reviewed it with the child before the probe session began. After each activity was completed, the child crossed out the picture and word representing the activity.

During the comparison phase, two interventions, *descriptive comments* (DC) and *descriptive comments plus trials* (DC+T), were used to teach letter names. There were some procedures that were the same across both conditions. These are described first, followed by the procedures that were specific to each intervention. A comparison of the components of each intervention is shown in Table 4.

Procedures identical across interventions. Before each session began, the child was presented with a choice board displaying four photographs of letter activities. Each option was reviewed, and the child was asked what game he would like to play. Choices were presented on the each child's board systematically. Prior to the first intervention session, the PI randomly selected four of the six possible letter activities and placed their corresponding photographs on the choice board. The field of four letter activities remained the same for the following intervention session. This means the same four letter activities were available for adjacent DC and DC+T sessions. To rotate the activities available to the child, after every other intervention session, activities the child had chosen for the previous two sessions were exchanged for those that had not been presented. If the same letter activity was chosen for both sessions in the adjacent pair (i.e., he chose to play fishing for letters in the DC and the DC+T sessions), only that activity was removed and was replaced with the activity that had not been recently played.

Table 4

Components of Interventions

Component	Interventions	
	DC	DC+T
Child chose the activity from four possible options.*	✓	✓
Adult sat in close proximity to the child.	✓	✓
Adult had positive verbal exchanges with the child during the session.*	✓	✓
Adult gave positive physical contact during the session (e.g., high fives, pat on the back).	✓	✓
Adult followed the child's lead during play.	✓	✓
Adult delivered a set number of descriptive comments during the session.*	✓ 10-12	✓ 2-6
Adult ensured the child was attending to the stimuli, either by following the child's lead or recruiting the child's attention, before delivering a descriptive comment.	✓	✓
Adult delivered four constant time delay instructional trials for each letter during the session.*		✓
Adult ensured the child was attending to the stimuli, either by following the child's lead or recruiting the child's attention, before delivering the task direction.		✓
Adult delivered praise when the child produced spontaneous or imitative targets.	✓	✓
Adult delivered praise after child responded correctly to instructional trials.*		✓

Note. DC = descriptive comments; DC+T = descriptive comments plus trials; *Indicates fidelity was collected on the implementation of that procedure.

All activities included pieces with target letters in the set being taught (i.e., set 1 letters for DC intervention, set 2 letters for DC+T intervention). No other letter stimuli were included in the activities.

The time of day the sessions took place was counterbalanced across adjacent pairs of interventions. For example, on the first day intervention sessions occurred, the DC intervention was used in the morning and the DC+T intervention was used in the afternoon. The following day, the order was reversed, and the DC+T intervention was used in the morning, and the DC intervention was used in the afternoon. If an intervention session did not take place, the sequence of interventions did not change. For example, if the child left early and an afternoon session was canceled, the intervention scheduled to be used that afternoon was used the following morning. At times, this resulted in the same intervention occurring in both morning and afternoon sessions, and counterbalancing on the predetermined schedule was not implemented as planned.

Each session was timed, and was targeted to last between 8 to 10 minutes. However, because priority was given to following the child's lead during intervention sessions, some sessions ran longer and some were shorter. For example, if the child began a play scheme during the session (i.e., building ice cube towers with the *Don't Break the Ice* pieces), the session might have been longer to complete the required components of the intervention (i.e., number of descriptive comments and/or trials). At the beginning of each session, the stopwatch or timer was started and a green go sign was shown to the child. The child was reminded that part of the game was looking at the letters and talking about them. A warning was given near the end of the session (e.g., "One more race and then it will be time for the red light."). At the end of the session, the go sign was flipped over to show the red stop sign. The child received a sticker for

participating in the game. All intervention sessions were audio taped and reviewed for data collection purposes.

Descriptive comments intervention. When the session began, the child was allowed to manipulate the materials as he chose. The PI demonstrated the use of new materials as needed (i.e., how to use the magnetic fishing pole to pick up fish) and occasionally suggested a new routine (i.e., "We could make cookies with the letters!"), but typically the child selected the theme of the play and directed the PI's role in the play. The PI was a play partner and maintained a back and forth exchange of comments and ideas throughout the session. The PI sat in close proximity to the child and often made positive physical contact with the child (e.g., gave a high five, pat on the back) and maintained a positive affect (e.g., smiling, laughing, making encouraging statements). During this highly responsive play time, the PI monitored the interactions with the materials for opportunities to deliver a descriptive comment. These opportunities arose in two ways: either the child was looking at letter stimuli (e.g., "You're cooking the letter B, like basketball"), or the PI recruited the child's attention to the letter stimuli (e.g., "Look! This letter K penguin is on the top of the iceberg!"). Whenever possible, the PI delivered a descriptive comment when the child was attending to letter stimuli during the ongoing play. However, there were not always sufficient opportunities to deliver the minimum number of descriptive comments (i.e., 10 for each letter) by following the child's lead. Thus, the PI recruited the child's attention to letter stimuli. However, the PI ensured the child was attending to the material prior to delivering the comment.

Descriptive comments always included two parts: the name of the letter and a descriptor. Descriptors included: a general locator (e.g., "Hey! There's a letter B on my fish!"), connection to a proper name (e.g., "It's a B, like in your friend Brian's name."), a connection to a salient

word (e.g., “It’s the letter F like football.”), or a motor model (e.g., “It’s a B, see it has a line, bump, bump.”). For each letter in the set being taught, the goal was to deliver between 10 and 12 descriptive comments per letter for a total of 20 - 24 descriptive comments. Descriptive comments were delivered as the opportunity arose; no pre-determined delivery schedule (e.g., one descriptive comment per min) was followed.

If the child said the target letter during the session, the PI praised the child, (e.g., “That’s right!”). If the maximum number of descriptive comments had not yet been delivered, the PI also paired the praise with a descriptive comment (e.g., “You got it! That’s the letter W, it goes down up down up!”). The PI attempted to deliver praise every time the child produced an imitative or spontaneous target, but this did not always occur.

***Descriptive comments plus trials intervention*.** Descriptive comments plus trials (DC+T) intervention sessions occurred in the same manner described above, with two exceptions: the PI made fewer descriptive comments about each target letter as compared to the number used in the DC intervention and delivered constant time delay (CTD) trials. In each DC+T intervention session, the PI delivered two to six descriptive comments and four CTD trials per letter.

For Jamal, Reggie, and Keandre, the PI delivered two descriptive comments in the DC+T intervention. One descriptive comment about the target letter was delivered before CTD trials for that letter began and one more was delivered after trials for that letter ended. For Mikey, the number of descriptive comments delivered in the DC+T intervention was increased to achieve a more balanced number of letter exposures across interventions (described below). The PI delivered two descriptive comments before CTD trials began, up to two more between trials, and two after trials ended, for a total of four to six descriptive comments per letter.

The CTD prompting procedure began by the PI securing the child's attention and delivering the task direction (i.e., "What letter is this?"). The PI waited a specified amount of time (0 s or 3 s) before delivering a controlling prompt ("Say B"). During the first three intervention sessions, the PI delivered the controlling prompt immediately after the task direction. In all subsequent sessions, the PI waited 3 s for a child response. If the child responded correctly, the PI gave a praise statement that included the letter name (e.g., "Very good, that's the letter B!"). If the child responded incorrectly or did not respond after 3 s, the PI delivered the controlling prompt. If the child responded correctly to the controlling prompt, the PI gave a praise statement that included the letter name. If the child again gave an incorrect answer or did not respond, the PI gave feedback and the correct answer (e.g., "No. It is a B") and moved on. Four trials for each letter were presented, resulting in a total of eight trials per session. Trials were delivered based on child attention; there was no specified inter-trial interval.

Because the PI said the letter name when delivering descriptive comments and instructional trials, the total number of exposures to the letter during the session was calculated. The total number of exposures varied depending on the child responses made during CTD trials. For example, if the child made four unprompted correct responses, only four exposures occurred (i.e., every time the PI praised the correct response). If the child made four prompted correct responses, eight exposures occurred (i.e., every time the PI delivered the controlling prompt and every time the PI praised the correct response). Thus, when two descriptive comments were delivered in the DC+T intervention, the total number of exposures for each letter could range between six (four during CTD trials plus two descriptive comments) and ten (eight during CTD trials plus two descriptive comments). For Mikey, the number of descriptive comments

delivered in the DC+T intervention was increased as described above to achieve a more balanced number of letter exposures across interventions.

The PI monitored the opportunities to deliver a trial in the same manner she monitored the opportunities to deliver a descriptive comment. Opportunities arose in the same two ways opportunities for descriptive comments arose: either the child was looking at a letter stimuli (e.g., "The green car won! What letter is this?"), or the PI recruited the child's attention to the letter stimuli (e.g., "Look! I'm going hit this ice cube right here. What letter is this?"). Whenever possible, the PI delivered a trial when the child was already attending to letter stimuli. However, these opportunities were not always sufficient to deliver the minimum number of trials (i.e., four for each letter). Thus, when necessary, the PI recruited the child's attention to letter stimuli before delivering a trial.

Generalization. Two types of generalization sessions were conducted: generalization across materials and generalization across adults. Generalization across materials sessions were conducted three times during baseline, and then once per week until the end of the study. During these generalization sessions, the PI asked the child to choose a book to read together. A total of five books were presented during baseline. A different book was chosen in each of the three baseline sessions. During the comparison phase, the three books selected during baseline were offered on a systematic rotating basis. For Reggie, a book that was not read in baseline was inadvertently offered as a choice in one session during the comparison phase. During the generalization sessions, the PI pointed to each target letter in the book and delivered the task direction, "What letter is this?" Three trials for each of the six target letters were delivered in random order resulting in a total of 18 trials per session. The PI engaged in typical book reading

interactions. No descriptive comments about target letters were made and no prompting occurred.

Generalization across adults sessions occurred twice during baseline and once after each set reached criterion. During generalization across adults sessions, the child's classroom teacher invited the child to play the *Alien Grab* game during center time. Activity pieces with all the target letters and three known letters were included in the game. Prior to the first session, the PI explained the procedures and practiced with the teacher the correct delivery of the task direction, prompts and feedback. During the activity with the child, the teacher delivered the task direction, "What letter is this?" three times for each target letter. The teacher praised correct responses and gave neutral responses after errors and no responses. The teacher gave encouragement throughout if needed to maintain attention. No descriptive comments about letters and no prompts were provided. The teacher received immediate verbal feedback from the PI about her correct implementation of probe sessions. The researcher recorded all child responses on the appropriate data sheet and collected procedural fidelity.

Maintenance. Maintenance probes were conducted intermittently up to 5 weeks after the comparison phase ended. Maintenance probes followed the same procedures as probes conducted during baseline and the comparison phase.

Procedural Fidelity

Procedural fidelity data were collected using two systems. First, a checklist was developed to measure the extent to which procedures were conducted as planned during all conditions (e.g., baseline, comparison, maintenance). Five procedural fidelity checklists were used, one for each session type: (a) probe sessions, (b) DC intervention sessions, (c) DC+T

intervention sessions, (d) generalization across materials sessions, and (e) generalization across adults sessions. Procedural fidelity items included in each checklist are summarized in Table 5. The presence or absence of each item was indicated by circling "yes" or "no". For each checklist, procedural fidelity was calculated by dividing the number of items circled as "yes" by the total number of checklist items.

Second, procedural fidelity data were collected on the extent to which the interventionist implemented the two interventions (i.e., DC, DC+T) correctly. This type of procedural fidelity was collected by recording the presence or absence of the required elements of each intervention (i.e., DC and DC+T) during each trial. Three different data sheets were used to collect procedural fidelity: (a) descriptive comments, (b) 0 s time delay trials, and (c) 3 s constant delay trials. Procedural fidelity was calculated by dividing the number of elements present in each intervention component by the total number of elements possible for the intervention. Procedural fidelity checklists and data sheets are included in Appendix C.

The PI completed the appropriate procedural fidelity checklist or data sheet after every session. In addition, a trained graduate student observed at least 25% of all sessions to assess the extent to which procedures occurred as planned. The PI and observer scored all procedural fidelity data during intervention sessions. However, both reviewed the audio taped intervention sessions to ensure they had scored the information accurately. In addition, atypical occurrences that could account for variations in the data were noted (e.g., it was the first session after a prolonged absence or the class had been on a field trip that day).

Table 5

Procedural Fidelity Checklist Items across Types of Sessions

		Intervention		Generalization	
Checklist Item	Probes	DC	DC+T	Materials	Adults
1. The adult and child were in a one-on- one setting.	✓	✓	✓	✓	✓
2. The child was given a choice of activities.		✓	✓		
3. The adult presented three cards /probe trials per target letter.	✓			✓	✓
4. The session did not have more than three disruptive interruptions or interference from peers.	✓	✓	✓	✓	✓
5. The session was conducted in a positive manner with less than two negative interactions between the adult and child.	✓	✓	✓	✓	✓
6. The session lasted __ minutes.	✓less than 5	✓ 8-10	✓ 8-10	✓ 8-10	✓ 8-10
7. The adult made the correct number of descriptive comments.	✓ 0	✓ 10-12 letter 1 ✓ 10-12 letter 2	✓ 2[a] letter 1 ✓ 2[a] letter 2	✓ 0	✓ 0
8. The adult included the correct number of prompted trials.	✓ 0	✓ 0 letter 1 ✓ 0 letter 2	✓ 4 letter 1 ✓ 4 letter 2	✓ 0	✓0
Total possible checklist items	7	9	9	8	8

Note. DC = descriptive comments; DC+T = descriptive comments plus trials.

[a] Procedures were changed to increase the number of descriptive comments to 4-6 per letter for Mikey.

CHAPTER III

RESULTS

Results are presented as follows. First, procedural fidelity data for probe sessions, generalization sessions, and intervention sessions are described. Second, IOA data are presented for the dependent measures and child behaviors during intervention sessions. Third, data from the DC and DC+T interventions are presented. Fourth, efficiency of the two interventions is compared in terms of sessions, exposures, and time to reach criterion. Fifth, generalization and maintenance of expressive letter naming data are presented. Finally, child target use during intervention sessions (i.e., spontaneous and imitated target productions during both interventions and child responses during CTD trials) are presented.

Procedural Fidelity

Procedural fidelity data were collected by trained graduate research assistants during probe sessions, generalization sessions, and intervention sessions. Probe and generalization session fidelity data are shown in Table 6. Procedural fidelity data for intervention sessions are shown in Table 7.

For Jamal, procedural fidelity was calculated for 40.8% (20 of 49) of his probe sessions, 41.7% (5 of 12) of his generalization across materials sessions, and 50% (2 of 4) of his generalization across adults sessions. For Keandre, procedural fidelity was calculated for 46.2% (18 of 39) of his probe sessions, 40% (4 of 10) of his generalization across materials sessions, and 50% (2 of 4) of his generalization across adults sessions. For Reggie, procedural fidelity was

Table 6

Mean (Range) Percentages of Procedural Fidelity in Probe Sessions

Child	Probe	Gen. Materials	Gen. Adults
Jamal	95 (85.7-100)	85 (62.5-100)	85.7*
Reggie	100	87.5 (75-100)	85.7*
Keandre	91.3 (71.4 – 100)	87.5 (75-100)	78.6 (71.4 – 85.7)
Mikey	100	91.7 (87.5-100)	92.9 (85.7-100)

*Note.**Indicates no range is reported as all sessions had the same procedural fidelity.

Table 7

Mean (Range) Percentages of Procedural Fidelity in Intervention Sessions

	Checklist Items		Descriptive Comments		CTD Trials	
Child	DC	DC+T	DC	DC+T	DC	DC+T
Jamal	84.2 (66.7-100)	88.9 (66.7-100)	99.5 (97.5-100)	100		97.9 (90.6-100)
Reggie	86.2 (66.7-100)	99 (96-100)	99.6 (95.7-100)	100		99 (96-100)
Keandre	82.2 (66.7-100)	89.7 (77.8-100)	100	99.3 (90-100)		99.6 (96.7-100)
Mikey	94.4 (88.9-100)	91.7 (88.9-100)	99.2 (97.4-100)	100		99.1 (92.9-100)

Note. DC = descriptive comments; DC+T = descriptive comments plus trials; CTD = constant time delay.

calculated for 43.5% (20 of 46) of his probe sessions, 41.7% (5 of 12) of his generalization across materials sessions, and 50% (2 of 4) of his generalization across adults sessions. For Mikey, procedural fidelity was calculated for 30 % (10 of 30) of his probe sessions, 42.9% (3 of 7) of his generalization across materials sessions, and 50% (2 of 4) of his generalization across adults sessions.

Procedural fidelity for probe sessions averaged over 90% for each child. The percent of checklist items marked "yes" was averaged across all sessions to obtain a mean percentage. Across all participants, the average procedural fidelity for probe sessions was 96.6% (range: 71.4 - 100%). The item most frequently marked no was "session lasted under 5 minutes." This item was met only 65% of the time in Jamal's probe sessions and 44% of the time in Keandre's probe sessions. Seven of Jamal's probe sessions and 10 of Keandre's probe sessions lasted longer than 5 minutes. All of the sessions that lasted longer than 5 min included the use of procedures to increase attending during probe sessions (i.e., a matching task, a scheduled break). Across all participants, the mean length of probe sessions was 4:57 (range: 0:56 - 20:04).

Procedural fidelity for generalization across materials was at an acceptable level. The percent of checklist items marked yes was averaged across all sessions to obtain a mean percentage. Across all participants, the average procedural fidelity for generalization across materials sessions was 87.9% (range: 62.5 - 100%). Again, the item most frequently marked no was related to session duration. The item "session lasted between 8 - 10 minutes" was met only 20% of the time for Jamal and Reggie, 25% of the time for Keandre, and 33% of the time for Mikey. This item was missed because many sessions lasted longer than 10 minutes. Across all participants, the mean length of generalization across materials sessions was 11:02 (range: 8:24 - 12:27).

Procedural fidelity for generalization across adults was also at an acceptable level. The percent of checklist items marked yes was averaged across all sessions to obtain a mean percentage. Across all participants, the average procedural fidelity for generalization across adults sessions was 85.7% (range: 71.4 - 100%). Again, the item most frequently missed was "session lasted between 8-10 minutes". This item was consistently missed because sessions lasted less than 8 minutes. Across all participants, the mean length of generalization across adults sessions was 6:17 (range: 5:11 - 9:38).

Procedural fidelity data were collected in 31.4% (11 of 35) of Jamal's DC intervention sessions and 50% (10 of 20) of his DC+T intervention sessions; 50% (10 of 20) of Keandre's DC intervention sessions and 35.9% (14 of 39) of his DC+T intervention sessions; 38.2% (13 of 34) of Reggie's DC intervention sessions and 28.6 % (4 of 14) of his DC+T intervention sessions; and 33.3% (6 of 18) of Mikey's DC intervention sessions and 32% (8 of 25) of his DC+T intervention sessions. Procedural fidelity data were collected for checklist items present and treatment integrity data were collected for the two interventions. Data for each participant and intervention (DC and DC+T) are shown in Table 7.

Across all children, procedural fidelity for checklist items for the DC intervention was 86.8% (range: 82.2 - 94.4%). Procedural fidelity for checklist items for the DC+T intervention was slightly higher than the DC intervention at 91.8% (range: 88.9 - 94.4%). Checklist items most often missed related to the number of descriptive comments delivered during the session. The target number of descriptive comments per DC session was 10 - 12 per letter. This item was only met 59.2% of the time. Generally, fewer comments were delivered than planned.

Across both interventions, descriptive comments were implemented with high treatment integrity. For the DC and DC+T interventions, mean treatment integrity for descriptive

comments was 99.6% (range: 99.2 - 100%) and 99.8% (range: 99.1 - 100%), respectively. Across all participants, the CTD procedure was implemented with high treatment integrity in the DC+T intervention, at 98.9% (range: 97.9 - 99.1%).

Interobserver Agreement

IOA data for probe and generalization sessions are shown in Table 8. For Jamal, IOA was calculated for 40.8% (20 of 49) of his probe sessions, 41.7% (5 of 12) of his generalization across materials sessions, and 50% (2 of 4) of his generalization across adults sessions. For Keandre, IOA was calculated for 46.2% (18 of 39) of his probe sessions, 40% (4 of 10) of his generalization across materials sessions, and 50% (2 of 4) of his generalization across adults sessions. For Reggie, IOA was calculated for 43.5% (20 of 46) of his probe sessions, 41.7% (5 of 12) of his generalization across materials sessions, and 75% (3 of 4) of his generalization across adults' sessions. For Mikey, IOA was calculated for 28.6 % (8 of 28) of his probe sessions, 42.9% (3 of 7) of his generalization across materials sessions, and 50% (2 of 4) of his generalization across adults sessions. The overall IOA across participants during probes was 98.7% (range: 83 - 100%). The overall IOA across participants during generalization across materials probes was 98.1% (range: 88.9 - 100%). The overall IOA across participants during generalization across adults probes was 96.2% (range: 72 - 100%).

IOA data for child behaviors during intervention sessions are shown in Table 9. IOA data were collected in 31.4% (11 of 35) of Jamal's DC intervention sessions and 50% (10 of 20) of his DC+T intervention sessions; 50% (10 of 20) of Keandre's DC intervention sessions and 35.9% (14 of 39) of his DC+T intervention sessions; 38.2% (13 of 34) of Reggie's DC intervention sessions and 28.6 % (4 of 14) of his DC+T intervention sessions; and 33.3% (6 of

Table 8

Mean (Range) Percentages of IOA in Probe Sessions

Child	Probe	Gen. Materials	Gen. Adults
Jamal	98.7(91.7-100)	97.8 (88.9-100)	86 (72 -100)
Reggie	98.3 (83-100)	98.9 (94.4-100)	100
Keandre	99.4 (88.9-100)	95.8 (88.9-100)	100
Mikey	98.3 (88.9-100)	100	100

Table 9

Mean (Range) Percentages of IOA in Intervention Sessions

	Imitative Targets		Spontaneous Targets		Responses in CTD Trials
Child	DC	DC+T	DC	DC+T	DC+T
Jamal	81.5 (0-100)	80 (0-100)	92.7 (66.7-100)	88.8 (50-100)	97.5 (75-100)
Reggie	89.7 (25-100)	33.3 (0-100)	79 (0-100)	74.1 (25-100)	79 (87.5 - 100)
Keandre	50 (0-100)	84.7 (0-100)	100	97.1 (60-100)	79 (87.5-100)
Mikey	69.5 (0-100)	75 (33.3-100)	88.3 (50-100)	100	96.8 (87.5-100)

Note. DC = descriptive comments; DC+T = descriptive comments plus trials; CTD = constant time delay.

18) of Mikey's DC intervention sessions and 32% (8 of 25) of his DC+T intervention sessions. Across all participants and all child behaviors, average IOA was 77.3% (range: 50 - 100%) for the DC intervention and 86.7% (range: 33.3 - 100%) for the DC+T intervention. Across participants, the mean IOA for imitative targets was 66% (range: 0 - 100%) for the DC intervention and 76.5% (range: 0 - 100%) for the DC+T intervention. Across participants, the mean IOA for spontaneous targets was 88.6% (range: 0 - 100%) for the DC intervention and 86.5% (range: 50 - 100%) for the DC+T intervention. Across participants, the mean IOA for child responses during CTD trials was 97.3% (range: 75 - 100%).

Acquisition of Expressive Letter Naming

Acquisition of expressive letter naming was measured by the percent of unprompted correct responses in probe sessions. Acquisition data are shown for each child in Figures 1 - 4. While the goal was to have one probe session and two intervention sessions each day, this did not always happen. Thus, the number of intervention sessions between probe sessions is not always the same. Variations occurred for several reasons. While children were available for study procedures during specific periods of time (i.e., between breakfast and morning recess), this time could vary based on the teachers' daily plans. For example, if a guest was coming into the classroom, the amount of time available for study procedures was reduced. Also, children were picked up at unpredictable times, which often meant the afternoon session could not be conducted. Because probes occurred each morning, it was not always possible to know, prior to conducting the probe, if the daily schedule was going to affect intervention sessions for that day. In addition, on some days, generalization across materials sessions occurred in place of an intervention session.

Jamal. Probe data for all three letter sets for Jamal are shown in Figure 1. During baseline, probes for expressive letter naming remained at 0% unprompted correct responding for sets 1 and 2. No change was seen until the intervention was applied.

During the comparison phase, when the DC intervention was applied to set 1, Jamal's performance remained at 0% until a change in level occurred following the fifth DC intervention session. Data rose to 50% unprompted correct responses but remained at that level because Jamal was consistently naming the letter D but not the letter F. Data remained at the same level until after the 13th DC intervention session when an increasing trend and change in level was observed. Data reached 100% for the first time following the 15th DC intervention session, and criterion was met following intervention session 17. Data remained stable at 100% until the comparison phase ended.

When the DC+T intervention was applied to set 2, an immediate increase to 50% was observed. Data remained stable at 50% for 18 intervention sessions. Jamal consistently named the letter K, but never named the letter G. Criterion was met following DC+T intervention session 20, and the comparison phase ended.

As shown in Table 10, Jamal reached criterion with the DC intervention (17 sessions) in fewer sessions than the DC+T intervention (20). Therefore, DC was considered the "superior" intervention and was applied to the control set (set 3) after the comparison phase ended. Following the third probe session during baseline, data for the control set rose to 50% because Jamal consistently named Q but not T. The control set data remained at 50% until the superior intervention (DC) was applied to set 3. After the 10th DC intervention session for set 3, an increasing trend was seen in the probe data. Data reached 100% for the first time after 14 DC

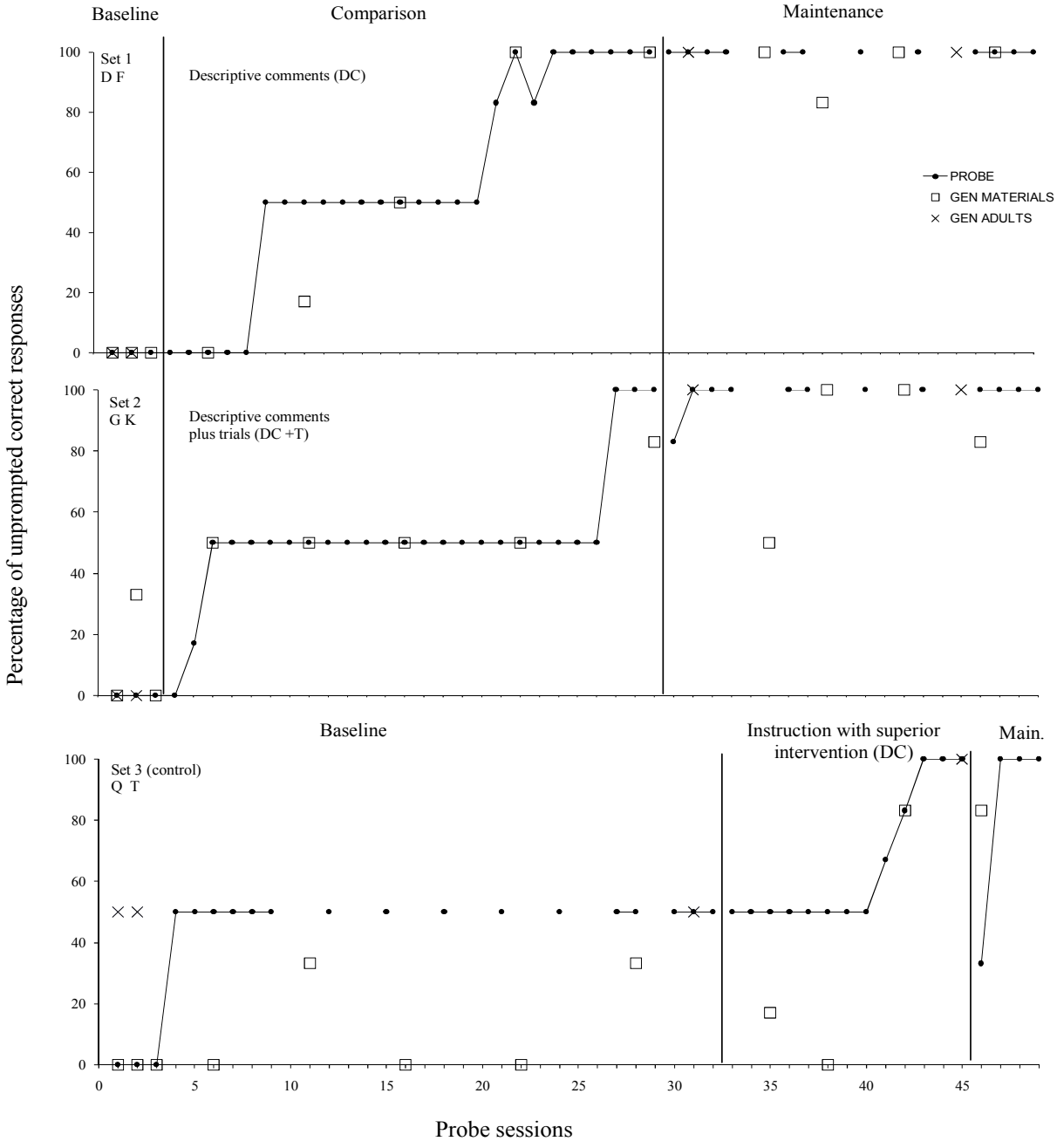

Figure 1. Probe session data for Jamal.

Table 10

Efficiency Data

	Sessions to Criterion			Average Exposures per Session			Exposures to Criterion			Average Length of Session			Minutes to criterion		
	Set 1	Set 2	Set 3	Set 1	Set 2	Set 3	Set 1	Set 2	Set 3	Set 1	Set 2	Set 3	Set 1	Set 2	Set 3
Child	DC	DC+T	DC	DC	DC+T	DC	DC	DC+T	DC	DC	DC+T	DC	DC	DC+T	DC
Jamal	17	20	16	22.8	14.1	21.1	388	281	338	11:18	10:27	10:53	192	209	174
Reggie	13	14	19	21	15.3	22.7	271	214	432	10:31	9:40	10:53	136	135	206
Keandre	*(20)	16	-	*(22.7)	15.8	-	*(454)	321	-	*(10:49)	9:34	-	*(216)	153	-
Mikey	14	*(24)	-	22.4	*(23)	-	314	*(553)	-	10:14	*(10:26)	-	143	*(250)	-

Note. DC = descriptive comments; DC+T = descriptive comments plus trials; * Indicates the set never reached criterion. Number reported in parenthesis is for all sessions because criterion was never met. Set 3 (control set) was only taught for Jamal and Reggie. Set 1 DC+T intervention for Keandre is not included in the table because it was not the initial intervention used for the set.

intervention sessions. Criterion was met after 16 DC intervention sessions, and the instruction with superior intervention ended.

Reggie. Probe data for all three letter sets for Reggie are shown in Figure 2. During baseline, probes for expressive letter naming in all three sets were stable with only one data point for each set above 0% correct responding. No changes were observed until the intervention was applied to each set.

Immediately following the application of the DC intervention to set 1, a slightly variable but increasing trend occurred. Data reached 100% for the first time following the sixth DC intervention session. Data were variable until criterion was met after 13 DC intervention sessions and then remained stable until the comparison phase ended.

When the DC+T intervention was applied to set 2, data remained at 0% correct responding until the probe session conducted after the seventh DC+T intervention session. Then, a slightly variable but increasing trend to 100% was seen. Data reached 100% for the first time following the 10th DC+T intervention session, but were variable until criterion was met after 14 intervention sessions. Once criterion was met, the comparison phase ended.

As shown in Table 10, Reggie reached criterion with the DC intervention (13) in slightly fewer sessions than the DC+T intervention (14). Therefore, DC was considered the superior intervention and was applied to the control set (set 3) after the comparison phase ended. Data for the control set was at 0% throughout the baseline and comparison phase with the exception of one data point. When the DC intervention was applied to the control set, there was a change in both trend and level. Data were variable until reaching 100% for the first time after 17 DC intervention sessions. Criterion was met after 19 DC intervention sessions, and instruction with the superior intervention ended.

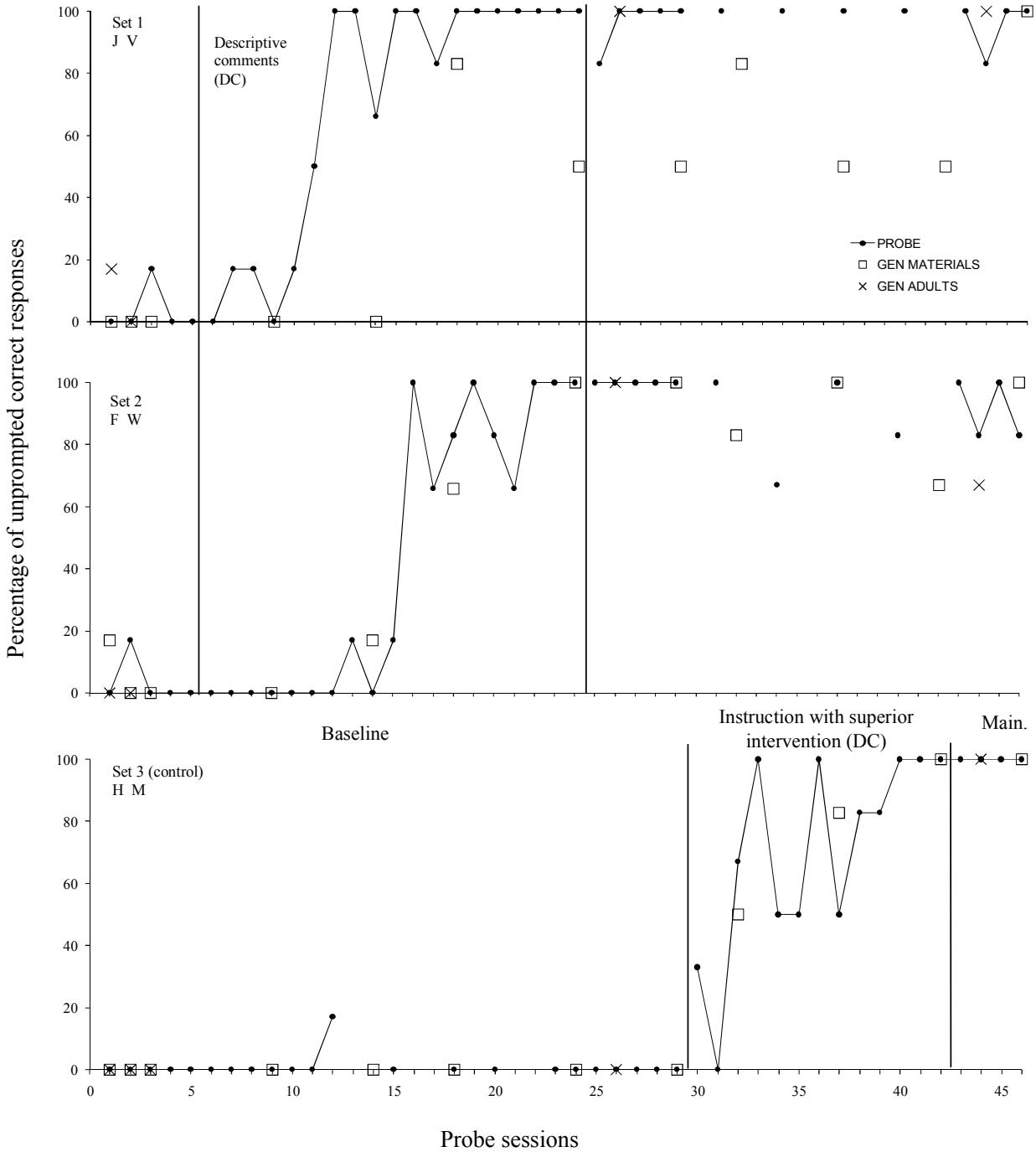

Figure 2. Probe session data for Reggie.

Keandre. Probe data for all three letter sets for Keandre are shown in Figure 3. During baseline, probes for expressive letter naming remained at or near 0% unprompted correct responses for all sets.

Throughout the comparison phase, correct responding for set 1 (DC) remained at 0%. After 20 DC intervention sessions on set 1 with no change in correct responding, the DC intervention was discontinued, and the DC+T intervention was applied to set 1. After 18 DC+T intervention sessions with no change in correct responding, the DC+T intervention was discontinued. When the DC+T intervention was applied to set 2, data remained at 0% correct until the probe session after the sixth DC+T intervention session. At that point, a variable but increasing trend to 100% correct responding was seen. Data for set 2 reached 100% unprompted correct for the first time following the 10th DC+T intervention session. Criterion for set 2 was met after 16 DC+T intervention sessions. At that point, the frequency of intervention sessions for set 2 was reduced to make time for more intervention sessions for set 1. From that point on, only two DC+T intervention sessions for set 2 were conducted. Data for set 2 remained stable for three probe sessions, and then variability increased and a downward trend was seen concurrent with a reduction of intervention sessions. During this period of time, there was also an increase in challenging behavior during probe sessions. Keandre would frequently put his head down on the table and refuse to participate. Because of these behavioral problems, probe sessions were discontinued for 2 weeks. When probes resumed, data for set 2 returned to 100%. Data remained at 100% for two more probe sessions, before falling to 67% in the final probe session.

The data for set 3 (control set) were at 0% correct responding throughout baseline and the comparison phase. Due to increasing frequency of challenging behavior, no intervention was applied to set 3.

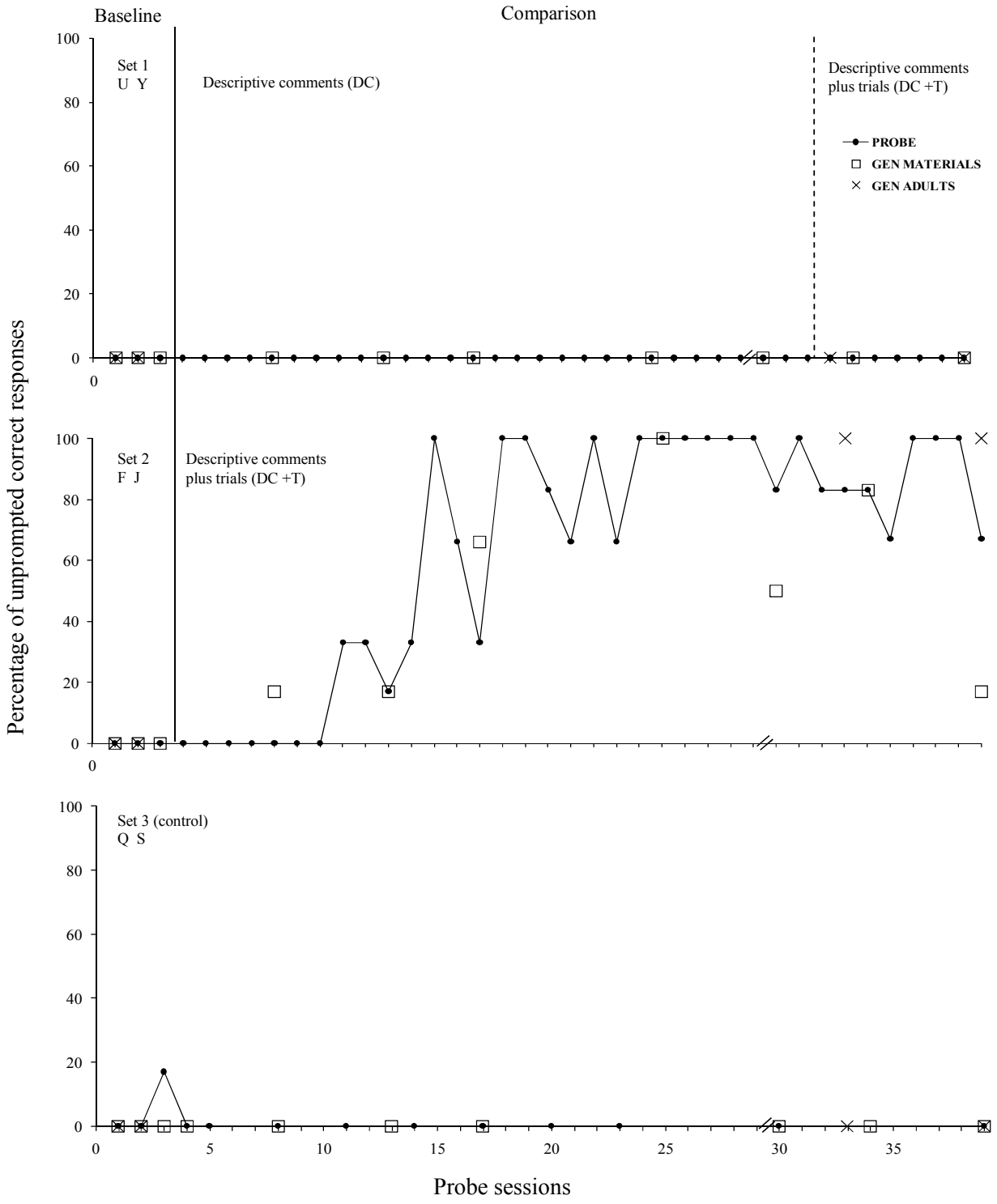

Figure 3. Probe session data for Keandre. Dashed lines on the abscissa indicate a one week child absence.

Mikey. Probe data for all three letter sets for Mikey are shown in Figure 4. During baseline, probes for expressive letter naming remained at 0% unprompted correct for all sets. No change in data was observed until the intervention was applied.

During the comparison phase, following the application of the DC intervention to set 1, data remained at 0% until a slightly variable but increasing trend began after the fifth DC intervention session. Data reached 100% for the first time following 12 DC intervention sessions, and criterion was met after 14 DC intervention sessions. Set 1 data remained stable at 100% until the comparison phase ended, with the exception of one missed trial during probe session 25.

When the DC+T intervention was applied to set 2, data remained at 0% until after the sixth DC+T intervention session. Data remained variable and ranged from 0-50%, with no apparent trend for 18 more intervention sessions. Intervention was discontinued after 24 DC+T intervention sessions and the comparison phase ended. Intervention continued past the a priori decision to end after 1.5 times the number of sessions to criterion for the other intervention because data appeared to have an upward trend. However, this trend did not continue. Due to time constraints, no intervention was applied to set 3.

Efficiency

The efficiency of the DC and DC+T interventions was compared using the data presented in Table 10. For each measure, data are compared only for the children who met criterion with both interventions (i.e., Jamal and Reggie). Then, the efficiency of the DC and DC+T interventions are compared across all stimulus sets that met criterion (n=8) including those sets for children who only met criterion in one set. Keandre's set 1 DC+T data were excluded

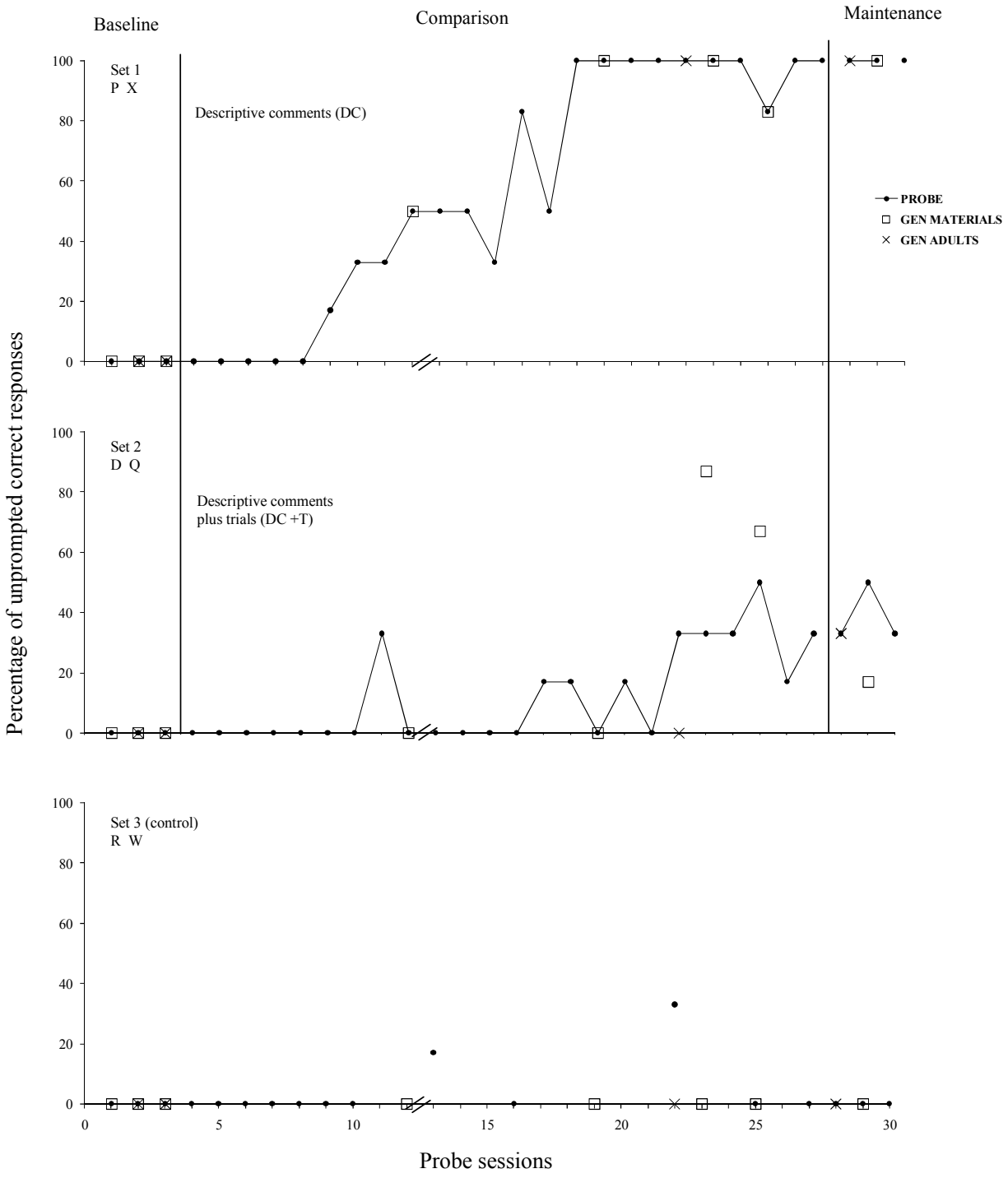

Figure 4. Probe session data for Mikey. Dashed lines on the abscissa indicate a three week child absence.

because the DC intervention had been applied to that set before the DC+T intervention was applied.

Three measures were used to compare the efficiency of the interventions: the number of intervention sessions until criterion was met, the number of exposures delivered during intervention sessions until criterion was met, and the number of minutes spent in intervention sessions until criterion was met. As previously described, the intervention with the fewest intervention sessions to criterion was considered "superior" when determining which intervention would be applied to the control set. However, two additional measures were considered related to efficiency. Data for each are presented below.

Sessions to criterion. For Jamal and Reggie, the DC intervention (set 1) required slightly fewer sessions to criterion than the DC+T (set 2) intervention. When the DC intervention was applied to Jamal and Reggie's control set (set 3), a slightly different pattern was observed across the two children. For Jamal, there was almost no difference in the number of sessions to criterion for set 1 (DC; *n*=17) and set 3 (DC; *n*=16), and both sets reached criterion in fewer sessions than set 2 (DC+T; *n*= 20). Reggie's set 1 and set 2 data were similar to Jamal's set 1 and set 2 data. There was almost no difference in the number of sessions to criterion for the two interventions during the comparison condition. However, when the DC intervention was applied to the control set (set 3), it took six more intervention sessions to reach criterion (*n*=19) than when the DC intervention was applied to set 1 (*n*=13). Thus, while DC was more efficient in terms of sessions to criterion than DC+T in the comparison phase, DC was less efficient (*n*=19) for set 3 than the DC+T intervention was for set 2 (*n*=14). Thus, it is not clear that the DC intervention was more efficient than the DC+T intervention for Reggie.

Across all children, criterion was met for 8 of the 10 letter sets that received instruction. Across those eight sets, five received the DC intervention and three received the DC+T intervention. When the DC intervention was used, the mean number of intervention sessions needed to reach criterion was 15.8 (range: 13 - 19). When the DC+T intervention was used, the mean number of intervention sessions needed to reach criterion was 16.7 (range: 14 - 20). Therefore, the DC intervention required an average of 0.9 fewer sessions to meet criterion than the DC+T intervention.

Exposures to criterion. Data on the average number of exposures per intervention session in each intervention as well as the total number of exposures to criterion are reported in Table 10. An exposure was defined as an adult utterance that included the target letter. For the DC intervention, this occurred when descriptive comments were made. For the DC+T intervention, this occurred when descriptive comments were made or during trials when the target letter was used during the controlling prompt or during feedback at the end of a trial. For Jamal, Keandre, and Reggie, the average number of exposures per session was greater in the DC intervention than in the DC+T intervention. For Mikey, the average number of exposures for the two interventions was relatively equal as the PI made an adjustment to ensure that the number of exposures was more consistent across the two interventions.

The number of exposures to criterion is a particularly relevant measure of efficiency given the discrepancy in the number of exposures per session for three of the children. Number of exposures to criterion will only be discussed for the two children who reached criterion in both interventions (i.e., Jamal and Reggie). For Jamal and Reggie, the DC intervention required more exposures to criterion than the DC+T intervention during the comparison phase. The same was true when the DC intervention was applied to Jamal and Reggie's control set (set 3). Jamal

needed more exposures (1.38 times as many) to reach criterion for set 1 (DC; *n*=388) than in set 2 (DC+T; *n*=281). Fewer exposures were needed to reach criterion for set 3 (DC; *n*=338) than for set 1 (DC; *n*=388). However, more exposures (1.2 times as many) were required to reach criterion for set 3 (DC; *n*=338) than for set 2 (DC+T; *n*=281). Like Jamal, Reggie needed more exposures (1.26 times as many) to reach criterion for set 1 (DC; *n*=271) than for set 2 (DC+T; *n*=214). However, Reggie needed more exposures (1.59 times as many) to reach criterion for set 3 (DC; *n*=432) than for set 1 (DC; *n*=271). He needed more exposures (2.02 times as many) to reach criterion for set 3 (DC; *n*=432) than for set 2 (DC+T; *n*=214).

Across all children, criterion was met for 8 of the 10 letter sets that received instruction. Across those eight sets, five received the DC intervention and three received the DC+T intervention. When the DC intervention was used, the mean number of exposures needed to reach criterion was 348.6 (range: 271 - 432). When the DC+T intervention was used, the mean number of exposures needed to reach criterion was 272 (range: 214 - 321). Therefore, the DC intervention required an average of 76.6 more exposures to meet criterion than the DC+T intervention.

Time to criterion. The average length of session and the number of minutes to criterion are shown in Table 10. Across all children but Mikey, the average length of the DC+T sessions was shorter than the average length of the DC sessions. However, the average length of the sessions varied by less than a minute for all children except for Keandre, who had significant behavior problems.

Data on minutes to criterion will only be reported for the two children who reached criterion in both interventions (i.e., Jamal and Reggie). There was variability in number of minutes to criterion for Jamal and Reggie. Jamal needed slightly more time to reach criterion in

set 2 (DC+T; 209 min) than in set 1 (DC; 192 min). While less time was needed for Jamal to reach criterion in set 3 (DC; 174 min) as compared to set 1 (DC; 192 min), it was still less time than needed to reach criterion for set 2 (DC+T; 209 min). Reggie's total time to criterion was near equal for both interventions, and only differed by 1 minute. However, a significant difference was seen for Reggie's set 3 (DC; 206 min) where 1.5 times as many minutes were required to reach criterion as compared to the minutes to criterion in either set 1 (DC; 136) or set 2 (DC+T; 135 min).

Across all children, criterion was met for 8 of the 10 letter sets that received instruction. Across those eight sets, five received the DC intervention and three received the DC+T intervention. When the DC intervention was used, the mean number of minutes to criterion was 170.2 (range: 136 - 206 min). When the DC+T intervention was used, the mean number of minutes to criterion was 165.3 (range: 135 - 209 min). Therefore, the sets taught using the DC intervention required an average of 4.9 more minutes to meet criterion than the sets taught using the DC+T intervention.

Generalization

Generalization across materials. Generalization across materials was assessed using letter stimuli included in a book. Generalization across materials data are shown in Figures 1 - 4. For Jamal (Figure 1) generalization across materials data were at 0% for all sets during baseline with the exception of one session for set 2 which was at 33%. As Jamal reached criterion in the DC intervention, his performance in generalization across materials sessions increased and remained high throughout the study. After the DC+T intervention was applied, an immediate increase to 50% was observed in the generalization across materials sessions. While there was

one data point over 50%, the data never reached 100% in the comparison condition. During the maintenance condition, generalization across materials data dropped back to 50% before increasing to 100%. Data then remained between 80 - 100%, similar to what was observed in his set 1 data. For set 3, generalization data ranged from 0 - 50% until instruction (DC) was applied. The highest generalization across materials data point for set 3 was at 80%.

For Reggie (Figure 2), generalization across materials data were at 0% for all sets during baseline with the exception of one session for set 1 (17%) and one session for set 2 (17%). When the DC intervention was applied to set 1, data remained at 0% for the next two generalization sessions. As Reggie began demonstrating the target letters during probes, an increase was also observed in the generalization sessions. However, data did not reach 100%, and returned to 50% before the comparison phase ended. During the maintenance phase, the majority of generalization sessions were at 50%, although data reached 100% during the last session of the phase. When the DC+T intervention was applied to set 2, generalization data increased following the eighth intervention session and reached 100% following 14 DC+T intervention sessions. During maintenance, data for set 2 (DC+T) were variable, ranging from 67% to 100% but were generally higher than the generalization data for set 1 (DC). Set 3 data remained at 0% until the DC intervention was applied. An increasing trend was observed until 100% was reached following the 11th intervention session. The same level was observed during the only maintenance session.

For Keandre (Figure 3), generalization across materials data were at 0% unprompted correct during baseline for all sets. There was no change in generalization data for either set 1 or set 3 throughout the study. For set 2, data were at 0% unprompted correct during baseline and gradually increased as the child began acquiring these letters in the DC+T intervention.

Generalization data first increased following the fourth DC+T intervention session. After reaching 100% during the seventh generalization across materials session, which corresponded to increases in the intervention data, a decreasing and variable trend back to near baseline levels was observed by the end of the study.

For Mikey (Figure 4), generalization across materials session data were at 0% for all sets during baseline. After the DC intervention was applied to set 1, generalization data increased as he began demonstrating the target letters during probe sessions. The first increase was seen after the seventh DC intervention session. Data remained above 80% until the end of the study. After the DC+T intervention was applied to set 2, data remained at 0% until an increase to 87% was observed after 17 DC+T intervention sessions. This was notable because it was a higher percent of unprompted correct responses than ever observed in probe sessions, and remained higher than probe sessions for the next generalization session. Generalization data did not maintain during the maintenance phase.

Because Jamal and Reggie met criterion with the DC and DC+T interventions, it was possible to compare the relative effectiveness of the interventions on generalization across materials. Generalization across materials occurred in all sets that reached criterion. Neither intervention showed superiority in generalization across materials settings. Stronger generalization outcomes were seen for Jamal in set 1 (DC) and for Reggie in set 2 (DC+T). Strong generalization outcomes were seen for both children in set 3 (DC).

Generalization across adults. Generalization across adults was measured by having the classroom teacher conduct probes with the target children. Data for generalization of expressive letter naming during play sessions with the classroom teacher are shown for Jamal, Reggie, Keandre, and Mikey in Figures 1, 2, 3, and 4, respectively. Similar data patterns were seen for all

sets and participants. During baseline, data were at 0% for all participants. The exception was Jamal's set 3 (control), which was at 50% because he named the letter Q and did not name the letter T. For each set, after intervention was applied and criterion was met for that set, another generalization across adults session occurred. Jamal, Keandre, and Mikey were able to name the stimuli for all sets that reached criterion at 100%. Reggie's data were slightly different. While his generalization data for sets 1 and 3 (DC) were all at 100%, his final generalization across adults probe for set 2 (DC+T) was at 63%. For all children, sets that did not reach criterion did not generalize across adults.

Maintenance

Maintenance probes were conducted either after criterion was reached for both sets (i.e., Jamal and Reggie) or after instruction ended (i.e., Mikey). With the exception of one data point, Jamal maintained 100% correct responses for set 1 (DC) and set 2 (DC+T) for 4 weeks after instruction ended. For Reggie, maintenance effects were clearly stronger in set 1 (DC). For set 1 (DC), data remained at or near 100% correct responding for 5 weeks after criterion was reached and instruction ended. For set 2 (DC+T), data remained at 100% for 1 week and then were variable ranging from 67% to 100%. Except for one data point, Jamal and Reggie's set 3 (DC) data maintained at 100% 1 week after criterion on set 3 was reached. Mikey also maintained 100% correct responding for 1 week in set 1 (DC) after instruction ended. No maintenance data were collected for Keandre.

Intervention Session Data

Child target production. Data for spontaneous, imitative, and CTD target productions during intervention sessions are shown in Table 11. To look at overall number of child productions that occurred across interventions, any prompted or unprompted correct letter naming during CTD trials were recorded as "target productions". In addition, productions that occurred in the DC intervention and outside CTD trials in the DC+T intervention were coded as either spontaneous or imitative. These were considered child-initiated productions as they did not follow a mand.

All children said their targets more per session in the DC+T intervention than in the DC intervention. Across children, an average of 1.66 (range: 1.28 – 2.12) times as many productions per session occurred in the DC+T intervention than in the DC intervention.

Target productions could be initiated by the child or the teacher. Spontaneous and imitative productions were considered child-initiated. Productions that occurred during CTD trials were considered teacher-initiated. The percent of child-initiated target productions was calculated by adding the number of spontaneous and imitative productions and dividing by the total number of productions. The percent of child-initiated targets made during the DC+T intervention were similar for Jamal (43.5%) and Reggie (44.3%). Mikey had a lower percentage of child-initiated target productions (32.5%) than Jamal and Reggie. Keandre had the lowest percentage of child initiated target productions (5.7%).

Differences across children were found between the types of child-initiated target productions that the children used. Across both interventions, Jamal and Reggie used more total spontaneous target productions than imitative target productions. The opposite was seen for

Table 11

Child Target Productions to Criterion

	Total Imitative		Total Spontaneous		Total During CTD	Total Productions		Average per Session	
Child	DC	DC+T	DC	DC+T	DC+T	DC	DC+T	DC	DC+T
Jamal	65	9	122	64	168	187	241	5.7	12.1
Reggie	91	18	152	33	115	243	166	7.6	11.9
Keandre	20	6	20	6	212	80	224	4	6.6
Mikey	77	48	38	14	191	115	253	8.2	10.5

Note. DC = descriptive comments; DC+T = descriptive comments plus trials; CTD= constant time delay.

Mikey, who made more imitative target productions than spontaneous target productions. Keandre had about equal imitative and spontaneous target productions.

Child responding during constant time delay trials. Data on child responding during CTD trials were collected for DC+T sessions and are shown in Figures 5 - 8. When 3 s delay trials were implemented, Jamal (Figure 5) began naming set 2 letters following the task direction almost immediately. Data remained high throughout the comparison condition ranging from 63.5% to 100%. His probe data met criterion even though he did not reach criterion level responding during intervention sessions. When 3 s delay trials were implemented, Reggie (Figure 6) showed an increasing trend in the percent of unprompted correct responses. Data remained high, but did not reach 100% until the last two intervention sessions, when probe data were nearing criterion. When 3 s delay trials were implemented, Mikey (Figure 7) showed variable levels of unprompted correct responding throughout the condition. He reached 100% correct responding during the 18th DC+T intervention session, but then a decreasing trend was observed until intervention was discontinued. This corresponded with variability in probe session data.

Figure 8 shows Keandre's responses during CTD trials for set 1, the set that received the DC+T instruction following the DC instruction, and set 2, the set that received only the DC+T intervention. For set 2 (which received the DC+T intervention first), there was an increasing trend over time in his unprompted correct responses. The first intervention session Keandre had 100% unprompted correct responses (intervention session 11) followed his first probe session at 100% unprompted correct (probe session 15). However, as it became more difficult to direct his attention to the letter stimuli during play, variability was observed in the intervention session data. This variability also was present in the data for Keandre's set 1. Due to PI error, two 3 s

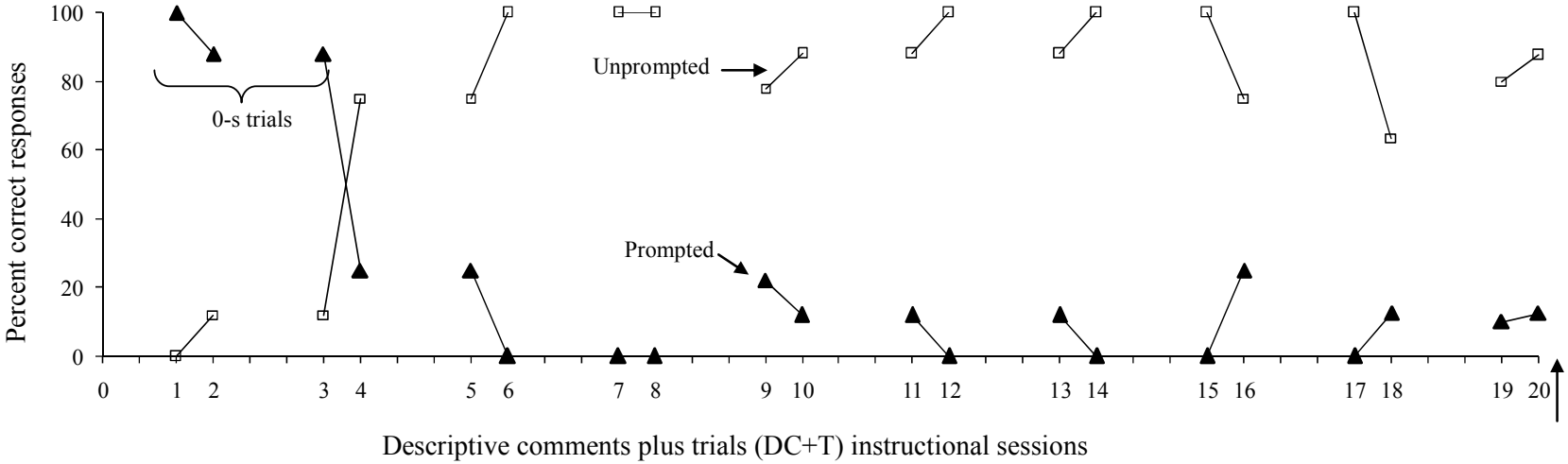

Figure 5. Percentage of prompted correct (filled triangle) and unprompted correct (unfilled square) responses during constant time delay for Jamal. Arrow indicates where criterion was met in probe sessions for set 2.

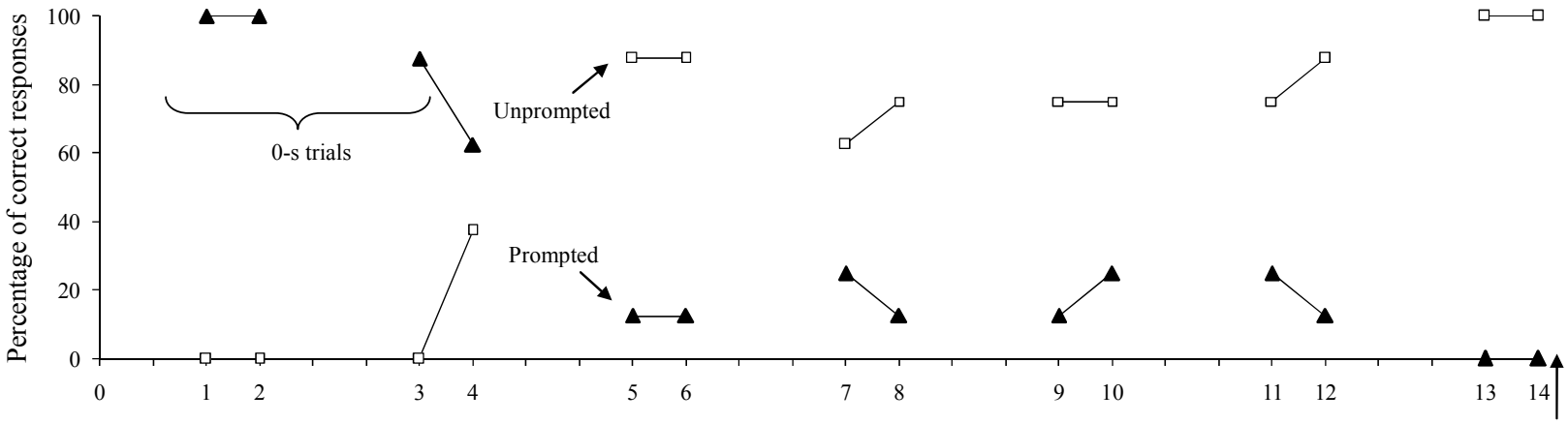

Figure 6. Percentage of prompted correct (filled triangle) and unprompted correct (unfilled square) responses during constant time delay for Reggie. Arrow indicates where criterion was met in probe sessions for set 2.

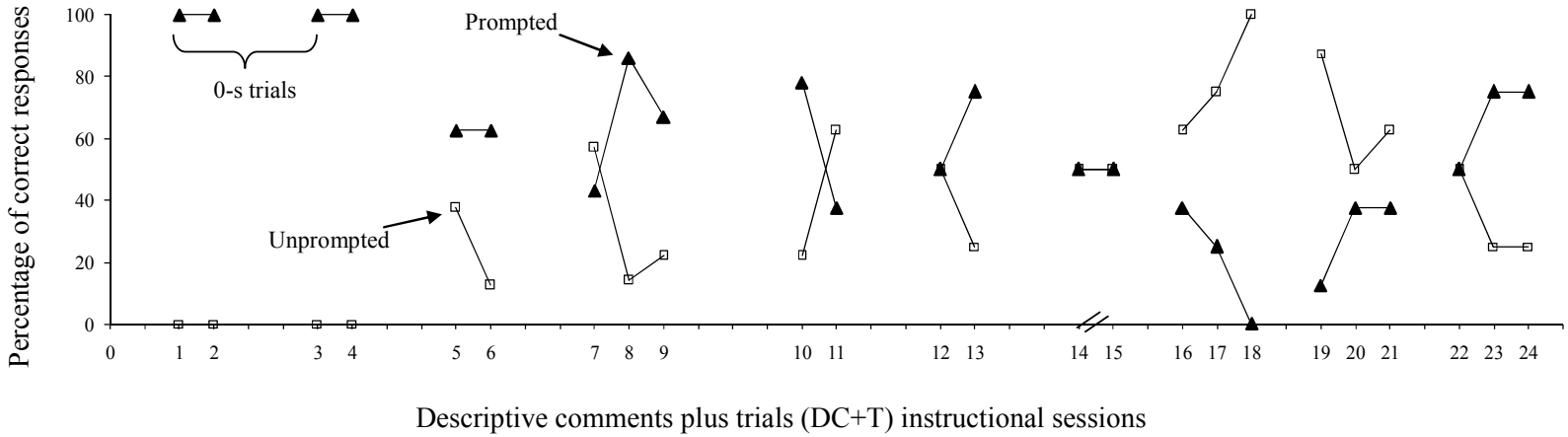

Figure 7. Percentage of prompted correct (filled triangle) and unprompted correct (unfilled square) responses during constant time delay for Mikey. Criterion was never met. Dashed lines on the abscissa indicate a three week child absence.

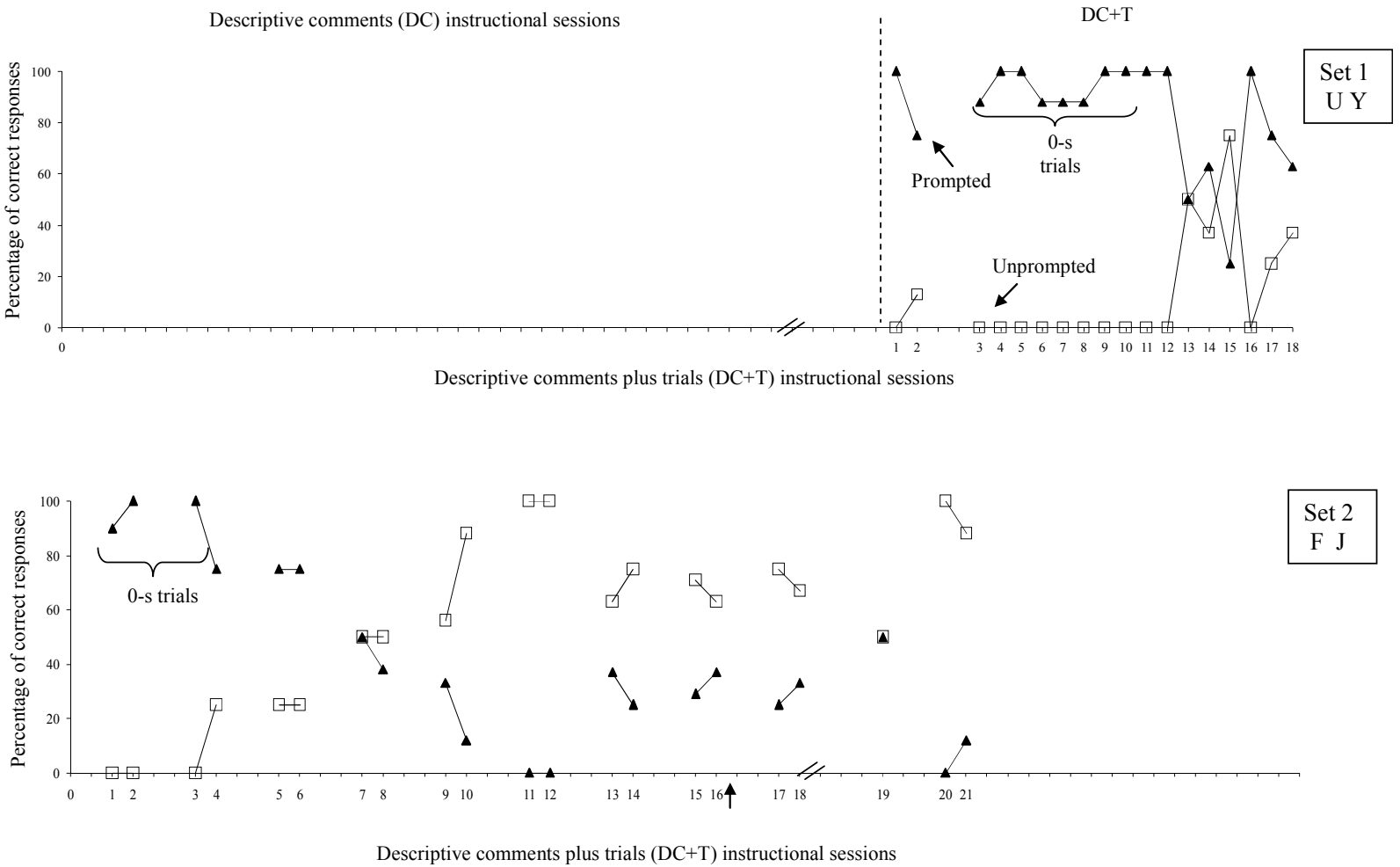

Figure 8. Percentage of prompted correct (filled triangle) and unprompted correct (unfilled square) responses during constant time delay for Keandre. Arrow indicates where criterion was met in probe sessions for set 2. Criterion was never met in set 1. Dashed lines on the abscissa indicate a one week child absence.

delay sessions were conducted before the 0 s trials began. At that point, 0 s delay sessions were conducted until he reached 100% prompted correct responding for three consecutive sessions. Nine sessions were required for Keandre to meet this criterion. During these sessions, although attention was secured at the beginning of the trial when the task direction was delivered, Keandre appeared to quickly withdraw his attention from the CTD trials, therefore failing to repeat the controlling prompt. Due to PI error, one additional 0 s delay session was conducted following the three 0 s delay sessions at 100%, and was also at 100% prompted correct. When 3 s delay trials began, unprompted correct responding increased to 50% for one session, but was variable for the remaining sessions. Intervention was discontinued after the 18th intervention session because challenging behavior continued to escalate and no increase in probe data was observed.

CHAPTER IV

DISCUSSION

The purpose of this study was to evaluate the relative effects of a commenting intervention (DC) and a commenting plus trials intervention (DC+T) for teaching letter naming during play activities with an adult. Both interventions included systematically delivered descriptive comments to expose children repeatedly to targeted letter names in a salient context. One intervention (DC+T) also included a CTD procedure. The effects of each intervention on the acquisition, maintenance, and generalization of letter naming were measured. In addition, three measures of efficiency were used to compare the interventions.

Overall, two children acquired, maintained, and generalized letter naming skills under both interventions. For the other two children, only one intervention was effective, and a different intervention was effective for each child. When the interventions were compared in terms of efficiency, differences were minimal. Across participants, the DC intervention required slightly fewer sessions but more letter exposures for children to reach criterion than the DC+T intervention. Total time to reach criterion, however, was equivalent across the two interventions. Generalization and maintenance data were similar across the two intervention conditions.

Intervention Effectiveness

Both the DC and the DC+T interventions were effective for teaching letter naming to two of the four children. One intervention was effective for teaching letter naming to the other two children, but the intervention that was effective differed for each of them. Criterion levels were

reached for all but one letter set taught using the DC intervention (5 out of 6). This is a promising finding in that it was the first demonstration in the literature of descriptive comments being used to teach letter naming. Similar commenting interventions have been used to teach vocabulary (Riches et al., 2005; Roberston & Ellis Weismer, 1999; Wilcox et al., 1991), but this study extended the findings of the effectiveness of commenting interventions to teach a different pre-academic skill. Criterion levels were reached for all but one of the letter sets taught using the DC+T intervention (3 out of 4). This supports the findings of other studies that have used CTD to teach pre-academic skills (Chiara et al., 1995; Daugherty et al., 2001; Grisham-Brown et al., 2000; Werts et al., 1992; Wolery, Ault, & Doyle, 1992; Wolery et al., 1998). However, this was the first time comments that included the target stimuli and a salient descriptor were included systematically in an intervention that used CTD.

Jamal, Reggie, and Mikey all reached criterion in set 1 (DC). They all willingly participated in study sessions, and required minimal attending prompts during intervention sessions. Keandre's results were different in several ways. While he reached criterion with the DC+T intervention, Keandre was the only participant who did not reach criterion under the DC intervention. This was true in spite of the fact that he received the most number of DC intervention sessions (n=20) used with any child. Also, he was the only child who received both interventions on the same letter set (set 1). Following 20 DC intervention sessions, Keandre received 18 DC+T intervention sessions. Although the application of the DC+T intervention on set 1 did not increase probe data, some correct responses were seen during CTD trials in the intervention sessions (Figure 8). It is likely that Keandre's history with failing to name the set 1 letters combined with behavior problems during probe sessions interfered with his performance, despite multiple attempts to increase compliance using various reinforcers (i.e., breaks with

novel toy play, computer time). Thus, his behavior problems and previous experience in probe sessions likely influenced the outcome of the application of DC+T to his set 1 letters.

Comparing Intervention Efficiency

Given that both interventions were effective for teaching letter naming to some children, efficiency was important to examine. Three measures of efficiency were examined: sessions to criterion, exposures to criterion, and minutes to criterion.

The DC intervention required slightly fewer sessions to criterion than the DC+T intervention. This was true when comparing data for children who met criterion with both interventions and when comparing across all children's sets that reached criterion. There were two opportunities for intra-subject replication of this finding when the DC intervention was applied to Jamal and Reggie's control set. However, while the finding replicated for Jamal's control set, the opposite was found for Reggie. When the DC intervention was applied to Reggie's control set, he actually needed more sessions to reach criterion than he needed for set 2 (DC+T). The intra-subject replications provide support for the effectiveness of the DC intervention, but not as clear support for the efficiency of the DC intervention.

This finding contrasts with previous research, which found target words taught with mand interventions met criterion in fewer sessions than those taught with a commenting intervention. In a study by Kouri (2005), participants met target word acquisition criteria with fewer sessions when a mand intervention was used. However, study procedures varied from those used in the current study. In the Kouri study, in the mand intervention, many more trials per letter were delivered (i.e., 15 - 20), and sessions were longer (i.e., 45 - 50 min). These differences in procedures could account for the discrepancy in the results in several ways. It is

possible commenting interventions delivered during play are more effective when delivered in a shorter period of time, like the current study, rather than across a longer intervention session. In addition, Kouri did not report the types of models that were delivered. It could be that the addition of descriptors to a letter name improved the outcomes of the DC intervention in the current study. It is also possible the relative efficiency of commenting and mand interventions varies depending on what the type of pre-academic skill is taught.

Another study, Heal et al. (2009), found a mand intervention was more efficient than a commenting intervention in terms of sessions to criterion. Heal et al. compared the efficiency of a commenting intervention and a mand intervention on the acquisition of two sets of Spanish words across three participants. In that study, although the number of sessions each intervention required to reach criterion was not reported, intervention sessions ceased when one set reached criterion or 90 sessions occurred. Of the six sets, three sets reached criterion when the mand intervention was used, no sets reached criterion when the commenting intervention was used, and three sets never reached criterion. Criterion was determined by the number of correct responses during intervention sessions. However, this comparison gave an inherent advantage to the mand intervention, because teacher-initiated learning opportunities were only delivered in the mand intervention, and children rarely labeled target words spontaneously in either intervention. In the current study, neither intervention had an inherent advantage because acquisition of letter naming was measured during probe sessions rather than intervention sessions for both interventions.

In the current study, although the DC intervention required fewer sessions to criterion than the DC+T intervention, it required more exposures to criterion. This was true when

comparing data for the two children who met criterion with both interventions and when comparing across all children's sets that reached criterion.

Previous studies that compared commenting and mand interventions did not measure the number of exposures with the same level of detail as the current study, and they did not compare exposures as a measure of efficiency. Two of the comparison studies did not provide this level of detail about the application of the interventions (Cole & Dale, 1986; Yoder et al., 1995). Kouri (2005) reported a higher mean use of models during the commenting intervention, but did not measure total exposures to criterion. Ellis Weismer et al. (1993) reported a mean number of exposures for a percentage of sessions, but did not measure total exposures to criterion. Heal et al. (2009) did not measure exposures. Thus, the current study adds to the literature by measuring the number of exposures children received during the interventions and using these data as a measure of efficiency.

Overall, in the current study, the DC intervention and the DC+T intervention required nearly equivalent amounts of time to reach criterion. The number of minutes to criterion was similar for Jamal and Reggie when set 1 (DC) and set 2 (DC+T) were compared, but differences were seen for both children's set 3 (DC) in terms of minutes to criterion. While studies of mand interventions have calculated number of minutes to criterion as a measure of efficiency (Doyle, Wolery, Gast, & Ault, 1990; Tekin & Kircaali-Iftar, 2002), the use of this measure when comparing commenting interventions and mand interventions has not been explored in the literature.

When all three measures of efficiency are taken into consideration, although slight differences were seen, neither intervention was clearly superior in the current study. The lack of a clearly superior intervention has been found in previous studies comparing commenting

interventions to mand interventions to teach pre-academic skills (Cole & Dale, 1986; Ellis Weismer, et al., 1993; Yoder et al., 1995). Although these studies described slight differences in outcomes across interventions, there were no clear findings that indicated a superior intervention. The current study supports these findings and extends it by reporting additional measures of efficiency. This study measured efficiency by sessions to criterion, exposures to criterion, and minutes to criterion. Previous comparison studies reported efficiency in terms of growth on developmental measures in developmental skills, diversity and frequency of target words acquired, and number of learning opportunities.

Generalization and Maintenance

Another overall finding of this study was that generalization and maintenance of letter naming were observed for both DC and DC+T interventions. There was evidence of both generalization and maintenance for all sets that met criterion in both interventions. However, for a number of reasons, there is no clear evidence that one intervention resulted in better maintenance or generalization outcomes. First, because only two children reached criterion in both interventions, only their data were relevant for comparing the effects of the interventions on generalization and maintenance. Second, no clear differences were present even for those two children. A slight advantage was seen for the DC intervention over the DC+T intervention in the generalization across adults data. Near equivalent generalization of expressive letter naming across materials was seen in the DC and DC+T interventions for one child, while better generalization across materials was seen in the DC+T intervention for the other child. A similar pattern was seen for maintenance, but the opposite intervention was favored. Near equivalent

maintenance of expressive letter naming was observed in the DC and DC+T interventions for one child, while better maintenance was observed in the DC intervention for the other.

Only one study that compared commenting interventions to mand interventions measured generalization (Kouri, 2005). No studies compared maintenance. Kouri reported generalization of target words across settings was stronger in the commenting intervention than the mand intervention. This was not found in the current study, because when there was difference, a slight advantage was seen in the DC+T intervention for generalization across materials and in the DC intervention for generalization across adults. While previous research on commenting interventions has reported generalization (Wilcox et al., 1991) and maintenance (Harris & Reichle, 2004; Riches et al., 2005) outcomes, the number of studies in which these types of measures have been collected is limited. Research on embedded mand interventions for pre-academic skills has included measures of generalization more frequently than research on commenting interventions (Bambara et al., 1988; Chiara et al.; Study 2 in Fox & Hanline, 1993; Studies 1 & 3 in Horn et al., 2000; Losardo & Bricker, 1994; Studies 1 & 2 in Peck et al., 1989; Werts et al., 1992; Wolery et al., 1993) and maintenance (Bambara et al., 1988; Chiara et al., 1995; Daugherty et al., 2001; Studies 1 & 2 in Fox & Hanline, 1993; Losardo & Bricker, 1994; Werts et al., 1992; Wolery et al., 1998; Wolery et al., 1993; Ziolkowski & Goldstein, 2008), generally with strong outcomes reported. The current study adds to the literature by measuring generalization (i.e., across adults and materials) and maintenance for both the commenting intervention and the mand intervention.

Individual Differences

Previous research has suggested the effectiveness and efficiency of interventions may be influenced by child characteristics, such as cognitive ability, willingness to participate, learning style, social abilities, or linguistic level (Yoder et. al., 1995; Cole & Dale, 1986; Ellis Weismer et al., 1993; Fey, 1986; Kouri, 2005; Sorensen & Fey, 1992; Wilcox et al., 1991). One of the goals of single subject research is to examine what works for whom, under what conditions (Horner et al., 2005; Wolery & Dunlap, 2001). In the current study, several variables were considered in terms of what might have influenced the outcomes. It does not appear that beginning knowledge of letters had an influence on which intervention reached criterion first. In fact, the child who knew the most letters in screening (i.e., Reggie) and the child who knew the fewest letters in screening (i.e., Mikey) had similar outcomes in the current study. All children scored below age level across developmental measures, so the influence of ability level on child outcomes cannot be addressed in the current study.

Although other children required occasional prompts to attend to letter stimuli during intervention sessions, Keandre was the only child who required high numbers of prompts. Keandre was also the only child who demonstrated challenging behaviors at a high rate during intervention sessions. He was also the only child for whom the DC+T intervention was more effective. One possible explanation might relate to Keandre's ability to attend to the stimuli during play. Joint attention has been found to be a necessary component of focused stimulation approaches to increase speech production (Ellis Weismer & Robertson, 2006). Perhaps a similar level of attention is required for a commenting intervention to be effective for expressive letter naming. Attention was not measured in the current study. Anecdotally, however, multiple redirections were required for Keandre to attend to the letter stimuli when either a descriptive

comment or a CTD task direction was delivered. However, the CTD procedures ensured that Keandre attended at least enough to repeat the controlling prompt (e.g., "Say J"). As time progressed, Keandre became increasingly more non-compliant and irritable when the PI attempted to deliver descriptive comments or task directions during play sessions, and avoided attending to the letters by closing his eyes, looking away, or removing the materials with letter stimuli from his play routines. The increasing frequency of attention problems and challenging behaviors offer a possible explanation why the DC+T intervention was not effective when applied to set 1, even though it was effective when applied to set 2.

Previous research has reported outcomes that vary based on individual subject characteristics, supporting the findings of the present study. Ellis Weismer et al. (1993) reported that a commenting intervention was more effective than a mand intervention for one participant, the mand intervention was more effective than the commenting intervention for another participant, and neither produced gains for a third participant. They speculated personality and learning style differences played a role in these outcomes. For example, the child who liked to initiate performed better during the commenting interventions. Also, Kouri (2005) reported large standard deviations on outcome measures, indicating individual differences were present in that study as well.

Use of Descriptive Comments

In the current study, the systematic delivery of descriptive comments during play activities that were designed to maintain child participation, interest, and engagement was effective in teaching expressive letter naming for 3 out of 4 children. This was the first study to describe systematically and measure the application of a commenting intervention that paired the

target stimuli with a salient descriptor to teach a pre-academic skill. These salient descriptors may have provided extra support to the children as they were learning letter names.

Anecdotally, during probe sessions, when children were presented with the stimuli, rather than saying the letter name, they often responded with a comment that correctly described the letter and that was similar to the comments they had heard during the intervention sessions (e.g., when shown a D and asked "What letter is this?", they would say "Like Daddy!"). This could indicate the descriptive comments provided a relevant association for the letter that was easier to retain than the letter name. However, Keandre often responded with a correct letter descriptor and an incorrect letter name. This occurred most often during intervention sessions for his set 1 letters. For example, when presented the letter U, Keandre would frequently respond, "F, it goes down and up like a smile". He was saying a correct descriptor of the letter being probed but then naming a different letter. Perhaps he was able to say a correct letter descriptor before he could name the correct letter because the descriptor was more salient. In the limited attention he gave to the letter stimuli, the letter descriptor, which had saliency, rather than the letter name, which was arbitrary, was retained. It is also possible the additional information could have interfered with learning the target information, which was the letter name.

Limitations

Several limitations were present in the study and must be taken into consideration when interpreting the findings. First, as with other single subject experimental research studies, the findings can only be generalized to children who meet the same inclusion criteria as the children in this study and when letter naming skills are the target behavior. Second, because Mikey's DC+T intervention procedures were changed to include more descriptive comments, it is not a

true replication of the procedures used with the other children. Even though Mikey received more descriptive comments during the DC+T intervention sessions than Jamal, Reggie, or Keandre (for set 2), he did not reach criterion in this intervention. This suggests that more descriptive comments during the DC+T intervention might not have made the intervention more effective. However, it is possible that the number of exposures per session might interact with child characteristics to produce differential outcomes. It is not known what the outcomes would have been if the exposures for Jamal and Reggie were more equivalent across the two interventions.

Third, procedural fidelity was slightly lower for the DC intervention than the DC+T intervention. Across all participants, the fidelity for the checklist item related to the adult giving the correct number of comments was considerably lower in the DC intervention (33.3%) than the DC+T intervention (98.1%). Because the nature of the play sessions was to maintain child engagement and attention, the adult was focused on being responsive to the child by following his lead and then expanding on play themes with the underlying intent to create opportunities to deliver descriptive comments or prompted trials. This resulted in two challenges that contributed to the frequent lack of adherence to the procedures. First, the number of available opportunities to insert descriptive comments or prompted trials depended on the play routines. For example, during the cooking with letters activity, if the child put all the letter pieces in the pan it was difficult to single one out to deliver a descriptive comment or prompted trial. The PI had to create opportunities to single out letters (e.g., “Put just one on my bowl so I can taste it.”), and this took time to incorporate in the play naturally. Second, the focus required to create these opportunities while maintaining child engagement often meant the targeted number of descriptive comments presented during sessions did not always occur as planned. Because the

targeted number of descriptive comments was higher in the DC intervention, this item was more likely to be missed, which contributed to the overall lower procedural fidelity seen in the DC intervention. While these factors were also an issue for delivering CTD trials, the fidelity for the checklist item related to the adult giving the correct number of prompted trials in the DC+T intervention (67.9%) did not impact the overall procedural fidelity score as drastically. Regardless, it is important to note that across each child's sessions, the number of descriptive comments and trials that were delivered fell within the desired range. Thus, these procedural fidelity results were likely not a factor in the findings of this study.

The overall procedural fidelity scores for both the DC and DC+T interventions were affected by the inconsistent length of the sessions. The issues described above also influenced variations in session time. If a child was engaged in a particular play routine that did not lend itself to incorporating descriptive comments or prompted trials, the PI allowed the play to continue. Throughout the session, the PI continually attempted to build alternate play routines to maximize opportunities for delivery of comments or trials. Despite the lower fidelity scores for this checklist item, the overall average time spent in all sessions across all children and sets fell within a 2 min range (9:40 - 11:18), which was the original targeted time span. Furthermore, the average time spent in all DC intervention sessions (10:41) and the average time for all DC+T intervention sessions (10:12) were nearly equivalent. This procedural fidelity issue was the same across interventions, and thus did not affect the comparison of the two interventions.

Fourth, a procedural error was made in how the CTD intervention was implemented for three of the children (i.e., Jamal, Reggie, and Keandre). The CTD procedure involves the use of three consecutive sessions of 0 s trials with 100% prompted correct responses before delay trials began. While at least three sessions of 0 s trials were delivered for each child, they were not

always at 100% prompted correct responding. For Jamal, three sessions of 0 s trials were delivered but 100% prompted correct responding occurred in only one of those three sessions. For Keandre (set 2) and Reggie, only two sessions at 100% prompted correct responding were conducted. In all cases, the number of unprompted correct responses increased immediately when the delay trials began, indicating the number of 0 s trials was likely sufficient. However, it is still a limitation, because procedural fidelity has been found to impact the effectiveness of the intervention procedures (Holcombe, Wolery, & Snyder, 1994). When Holcombe et al. compared high and low implementation of procedural fidelity, significantly more errors were made (i.e., failure to deliver the controlling prompt in 50% of the sessions) than occurred in the current study. Regardless, it cannot be known if the errors influenced the variability in the intervention data or the number of sessions to criterion in the current study.

Fifth, these interventions were applied during highly engaging interactions with an adult who was skilled at creating unique play scenarios with creative materials, managing behaviors, and maintaining child attention. It is not known if these interventions would work in a less engaging interaction. Also, no procedural fidelity data were collected on these types of interaction variables, so it is not possible to determine if this was consistent across interventions or if it had any effect on the outcomes of the study.

Sixth, because the interventions were delivered by the PI in a one-on-one setting outside the classroom, it is not known if teachers would be able to implement DC or DC+T interventions in a classroom setting during ongoing activities and routines. One concern is whether or not teachers would be able to deliver consistently a high dosage of descriptive comments or CTD trials in the context of a busy classroom. Previous studies have shown that preschool teachers can be trained to deliver CTD trials to teach pre-academic skills during classroom activities and

routines (Grisham-Brown et al., 2000; Werts et al., 1992; Wolery, Anthony, Caldwell, Snyder, & Morgante, 2002). Compared to the substantial CTD body of literature, fewer studies on delivery of commenting interventions have been conducted in classroom settings (Cole & Dale, 1986; Wilcox et al., 1991; Yoder et al., 1995). No studies have combined CTD with descriptive commenting. Given the focus required for the PI to deliver the interventions in the current study, this is a particularly important issue to address in future research.

Seventh, a confounding variable is present for Jamal because he began naming one letter (Q) in his control set (set 3) immediately following the beginning of intervention for sets 1 and 2. He also named the letter Q during baseline in generalization across adults sessions, but not in generalization across materials sessions or probe sessions. While data for the other control stimulus (T) did not increase until the intervention was applied, this confound influences the extent to which experimental control was achieved.

Eighth, despite the extensive efforts to ensure equal difficulty across sets of stimuli, it is possible the letters selected for Keandre's set 1 (DC) were inherently more difficult than the letters selected for Keandre's set 2 (DC+T). He reached criterion on his set 2 target letters following 16 sessions of the DC+T intervention. However, he received both the DC intervention (20 sessions) and the DC+T intervention (18 sessions) on his set 1 letters and never reached criterion. While lack of equivalence in difficulty of letters might be one explanation (i.e., the set 1 letters were more difficult), it is also possible his challenging behaviors interfered with his learning. Keandre showed increasing challenging behavior during the DC intervention for set 1 which may have affected the extent to which the DC+T intervention was effective given that it was applied following 20 sessions of the DC intervention on the same set. It is not possible to

know what the effects of the DC+T intervention on set 1 would have been had it been applied first.

Future Research

This study extended the literature by applying the use of an intervention that included descriptive comments to a pre-academic skill other than vocabulary. In addition, the systematic implementation of descriptive comments during CTD trials extends the literature on the CTD procedure. This study also provided a comparison of a descriptive commenting intervention with a descriptive commenting intervention that also included mands. Further studies are warranted to more completely examine the effectiveness and efficiency of these two interventions. Possible directions for future research include changing the way components of the interventions are delivered, applying the same intervention in different settings or with different participants, or modifying procedures to address the limitations of the current study.

Changing the way intervention components were used in the current study could provide more information about the relative effects of each component. For example, one component of the intervention was the number of exposures delivered during the intervention. Given the finding that the DC intervention required fewer sessions to criterion but more exposures, future research might keep the number of exposures per session constant across interventions. It is possible that if the number of exposures had been kept constant across conditions, the data on number of sessions to criterion might have been different. For Reggie and Jamal, their number of sessions to criterion was fairly similar across interventions, but their number of exposures to criterion was significantly different across the two interventions. This suggests the need for

keeping the number of exposures per session more consistent to understand whether or not the number of exposures or number of sessions to criterion is an issue.

Another component of the interventions used in the current study that could be examined in future research is the feedback that was provided to the children. In the current study, children were praised when a spontaneous or imitative target occurred. Because the number of spontaneous and imitative target productions varied across children, this led to different amounts of praise delivered across sessions and across children. For example, Keandre, who made significantly fewer target productions during the intervention sessions, likely received much less praise than the other children. Because the delivery of praise was not coded, it is not known whether this influenced the findings. A future study could measure praise more precisely, eliminate the use of praise all together, or compare different amounts of praise to investigate the influence of praise on the outcomes of the interventions.

Another component that could be examined in future research relates to how the trials and comments were initiated. In the current study, some trials and comments were delivered based on what the child was attending to and others were delivered after recruiting the child's attention. When possible, the child's attention to the letter stimuli during play was the antecedent for delivery. However, when this did not result in enough opportunities to deliver the correct number of comments or trials, the child's attention was recruited. Data were not collected on the number of comments or trials that were delivered under each of these conditions, so it is not possible to determine if this affected the relative efficiency or effectiveness of the interventions. Future research might compare these two antecedent conditions to determine how they impact the effectiveness and efficiency of the interventions.

Another variation in the interventions that could be examined is the content and type of descriptor that was used as part of the comments. In the current study, the children often named the descriptor (e.g., "like Daddy") before they named the letter. It is not known if the descriptors help or hinder learning letter names. It is possible the descriptors increased the cognitive load, making it more difficult for children to recall the target information (i.e., the letter name). In the current study, a descriptor was a component of every descriptive comment. A future study should compare the use of comments with the letter name only to comments with the letter name and a descriptor. In addition, various forms of descriptors could be used to determine if a certain type of descriptor is more effective. These types of comparisons would help to better understand the function of providing descriptors when teaching letter names.

Future research should also examine the use of the interventions to teach different skills or to teach different types of children. For example, the use of descriptive comments and descriptive comments plus trials could be applied when teaching other pre-academic skills, such as naming shapes, numbers, or colors. To explore individual child differences, the same interventions could be used with children with different learning characteristics, such as children with speech language impairments or with other developmental disabilities. In addition, approaches to learning, ability to attend, and other child characteristics could be carefully measured and compared.

To understand the effects of these interventions in different contexts, future research should examine the interventions when implemented in ongoing classroom activities and routines. The goal would be for teachers to use these interventions in their everyday interactions with children in the classroom. Factors that influence the effectiveness of DC or DC+T interventions may change when delivered in a classroom setting. Finding enough opportunities

during the day to provide a similar concentration of exposures, adapting the environment to include enough letter stimuli, or finding adequate opportunities to embed comments and trials into engaging, child-directed play may present challenges that were not examined in the current study.

The current study was implemented by a skilled teacher. The engaging activities and interactive style of the teacher were components of both interventions in the current study. Future research should examine if teachers can be trained to not only learn the components of the intervention, but to deliver the intervention in the same manner as the PI. The skills used were complex (i.e., developing a relationship, creating unique play routines, embedding descriptive comments or trials), and previous research has shown it can be difficult to train teachers to use similar interventions (Yoder et al., 1995). While research exists on training teachers to use components of the interventions (e.g., CTD trials, responsive interactions, praise; Barton & Wolery, 2007; Kaiser, Ostrosky, & Alpert, 1993; McBride & Schwartz, 2003; Wolery et al., 2002), limited research exists on how to train teachers to use interventions with multiple components. However, there is a growing literature on the types of professional development strategies that result in teachers' use of interventions (Snyder, Artman, Hemmeter, Kinder, & Pasia, 2010). Future research on interventions such as those used in this study could investigate the use of these strategies for training teachers to use the interventions. Further, more consistent procedural fidelity data will be needed to understand if teachers can be trained to implement these interventions with fidelity in the context of ongoing activities and routines in the classroom.

Finally, future research should address limitations of the current study related to the use of probe sessions with the CTD procedure. In the current study, daily probes were used to collect

data during both intervention conditions in order to compare the two interventions. In previous studies of CTD, data were collected during the instructional trials (Wolery et al., 1992). It is possible that the use of daily probes in the current study affected child performance relative to the CTD procedure. During instruction with CTD, children are instructed to "wait if you don't know". This may have been confusing in the current study because the child received the controlling prompt during intervention sessions but did not receive any assistance during probe sessions. This issue could be addressed in future research by using intermittent probes or using intervention session data for the DC+T intervention. This would make the procedures more similar to those used in previous research on CTD. Another option would use an instructional procedure such as simultaneous prompting (Wolery, Ault, & Doyle, 1992) that requires the use of probe sessions for data collection.

Implications for Practice

A main finding of this study is that both interventions were effective for some children. In this study, some children learned to name letters when comments were delivered in the absence of mands and some children learned to name letters when comments were delivered in addition to mands. This may be valuable information in practice because teachers could tailor the intervention to the needs of the child. Teachers could take into consideration child characteristics that may make the delivery of a mand during play more or less desirable. For example, children who tend to be reserved, reluctant to participate, English language learners, or children from certain cultures where demands for production are not part of the typical conversational interaction between child and adult may benefit from commenting interventions that do not include mands (DeThorne et al., 2009; Ellis Weismer & Robertson, 2005). Reggie's teacher

reported at the beginning of the study he was unsure of his abilities and lacked confidence. It is possible by hearing many exposures in an engaging, enjoyable activity without pressure to respond he developed a positive learning history he had not previously experienced. Anecdotally, his mother described Reggie as "on fire" for letters after starting the study, asking her to tell him what other letters were that he saw in his environment. Prior to the study, Jamal's teacher reported he would often cry and refuse to participate when presented with a challenging task. It is possible the DC intervention was more successful for Jamal because he was not required to respond to task directions and received praise when he correctly named or imitated a target letter.

The findings from this study suggest mand interventions may be more appropriate for children who have difficulty attending and need more support to focus on stimuli presented during play. It is possible children like Keandre benefit from the brief shared attention required to respond when given a task direction or a controlling prompt to produce the target. In addition, children who have difficulty attending may not be able to discern the important information included in descriptive comments (i.e., the letter name) because their focus is on their play instead of the stimuli. Teachers should recognize child learning characteristics and find a match for the type of intervention which might work best given knowledge of the child's learning characteristics and history.

The effectiveness of the commenting intervention does not suggest that just talking about letters during play will result in children acquiring letter names. The level of environmental arrangement required and the interaction between the PI and child in the current study were intense. The materials and play routines were adapted to incorporate letters, and children were given frequent reminders that part of the game was looking at the letters and talking about them. It may be the case that this level of preparation and type of interaction is more difficult in a busy

classroom setting. The challenges presented in following the child's lead while providing opportunities for instruction have been discussed in the literature on naturalistic interventions (Hancock & Kaiser, 2005; Yoder et al., 1995; Yoder, Kaiser & Alpert, 1991).

While this study provides initial evidence to suggest providing multiple exposures to descriptive comments about letters can result in increased letter naming skills, teachers should be aware that some children, particularly those with attention problems, are likely to also benefit from the use of a prompting procedure such as CTD to support their learning. Alphabet knowledge is a highly valued skill, and making sure all children receive the support needed to learn letters in the most efficient way possible should be a goal of preschool programs. Further research on both the effectiveness and efficiency of interventions and matching the intervention to the characteristics of the child can expand the knowledge base related to effective and efficient instruction.

Appendix A

Teacher Literacy Surveys

Dear Teacher,

As part of your participation in the research study you are being asked to share information about the practices you typically use to support children's literacy learning in your classroom. Your responses will be confidential, and will NOT be shared with anyone. Thank you for your participation!

1. What kind of literacy curriculum or approach, if any, do you use? Some examples of literacy curricula are Read, Play, Learn; Letter People; Opening Worlds of Learning (OWL). Some examples of literacy approaches are letter of the week or whole language.

2. Please list <u>everything</u> you do that helps the children in your classroom learn about letters. What kinds of activities do you do, materials do you provide, strategies do you use…?

3.On the list you made above, please put a star next to the three that you use most often.

4.How much time, on average, do the children in your class spend each day in the following literacy related activities?

	Less than 15 minutes	15-30 minutes	30-45 minutes	More than 45 minutes
Listening to books being read				
Reading/ looking at books on their own				
Completing literacy related worksheets (about letters, rhyming, etc)				
Participating in a large group literacy activity other than book reading				
Participating in a small group literacy activity				
Participating in a literacy activity one on one with an adult				

5. How do you think children learn letters and other literacy related skills best?

Child's ID#:_____ Birthdate:_____

1. Gender:

___ Male

___ Female

2. Ethnicity (check all that apply):
___ African-American

___ Asian

___ Caucasian

___ Hispanic

___ Other (please specify: _____)

3. Does the child have any IEPs? (check one)

___ No (go to question #5)

___Yes (please describe:)_____

4. *(answer this question only if the child has an identified disability)*
What services does the child currently receive?
_____

How often? _____

Are any of the goals language or literacy related? _____

5. How long has this child attended this program? _____

6. What are this child's favorite activities?

(please turn over)

7a. Does this child usually stay engaged when: (check the best response for each)

	YES, (usually stays engaged)	**SOME-TIMES**	**NO** (often has trouble staying engaged)	**N / A** (does not happen in the classroom)
Listening to books being read?				
Reading/ looking at books on his or her own?				
Completing literacy related worksheets (about letters, rhyming, etc)?				
Participating in a large group literacy activity other than book reading?				
Participating in a small group literacy activity?				
Participating in a literacy activity one on one with an adult?				
Participating in center time play with other children?				
Participating in center time play independently?				

7b. Does this child have challenging behaviors that prevent him or her from being engaged in instructional activities?

_____YES

_____ NO

If so, indicate how often and please describe below. (check one)
___ occasionally___ 1-2 times per week ___1-2 times per day___ more than 2 times per day

Please describe the kinds of challenging behaviors you see:

_____

9. Is there anything else you would like to share about this child's literacy learning (specific struggles, strategies that work, etc)?

_____

_____

Appendix B
Data collection sheets

DATA SHEET –*Probes*

SESSION INFORMATION

Child ID #: _____ Date: _____ Session #: _____

Day of the week (circle) :M Tu W Th F Researcher initials: _____

Phase (circle): Baseline Intervention Maintenance Start Time:_____

IOA? ☐yes ☐no Stop Time: _____

CHILD RESPONSE DATA

circle the child response

TARGET LETTER (sets 1, 2, 3)	***TRIAL 1***	***TRIAL 2***	***TRIAL 3***
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR

Any behavioral observations?

DATA SHEET –*Descriptive Comments (DC)intervention session (SET 1)*

SESSION INFORMATION

Child ID #: _____ Date: _____ Session #: _____

Day of the week (circle) :M Tu W Th F Researcher initials: _____

Activity: _____ Start Time:_____

Session position (circle): first second Stop Time: _____

IOA? ☐yes ☐no

CHILD RESPONSE DATA

TARGET LETTER (set 1)	***How many comments about the letters did the child make during the session? (use tally marks)***	
	Spontaneous	Imitative
	Spontaneous	Imitative

Any behavioral observations?

DATA SHEET –*Descriptive comments plus trials (DC+T)intervention session (SET 2)*

SESSION INFORMATION

Child ID #: _____ Date: _____ Session #: _____

Day of the week (circle) :M Tu W Th F Researcher initials: _____

Activity: _____ Start Time:_____

Session position (circle):first second Stop Time: _____

IOA? ☐yes ☐no

CHILD RESPONSE DATA

TARGET LETTER (set 2)	circle the child response			
	TRIAL1	***TRIAL2***	***TRIAL 3***	***TRIAL 4***
	UC PC UE PE NR	UC PC UE PE NR	UC PC UE PE NR	UC PC UE PE NR
	Spontaneous comments (tally)		Imitative comments (tally)	
	UC PC UE PE NR	UC PC UE PE NR	UC PC UE PE NR	UC PC UE PE NR
	Spontaneous comments (tally)		Imitative comments (tally)	
Any behavioral observations?				

DATA SHEET - *Generalization across materials*

SESSION INFORMATION

Child ID #: _____ Date: _____ Session #: _____

Day of the week (circle) :MTuW ThF Researcher initials: _____

Book: _____

Phase (circle): Baseline Comparison Maintenance Start Time:_____

IOA?☐yes☐no Stop Time: _____

CHILD RESPONSE DATA

circle the child response

TARGET LETTER (sets 1, 2, 3)	***TRIAL 1***	***TRIAL 2***	***TRIAL 3***
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR

Any behavioral observations?

DATA SHEET - ***Generalization across adults***

SESSION INFORMATION

Child ID #: _____ Date: _____ Session #: _____

Day of the week (circle) :MTuW ThF Researcher initials: _____

Activity: __Alien Grab_____ Start Time:_____

Phase (circle):BaselineMaintenance Stop Time: _____

IOA?☐yes☐no

CHILD RESPONSE DATA

circle the child response

TARGET LETTER (sets 1, 2, 3)	***TRIAL 1***	***TRIAL 2***	***TRIAL 3***
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR
	UC UE NR	UC UE NR	UC UE NR

Any behavioral observations?

Appendix C
Procedural Fidelity

PROCEDURAL FIDELITY – PROBES	
SET UP / IMPLEMENTATION	(circle one)
Were the adult and child in a **one on one** setting?	YES NO
Did the adult present **3 cards per target letter**?	YES NO
Did the session have no more than three **disruptive interruptions / interference** from peers? *Describe what happened:*	YES NO
Was the session conducted in a **positive manner**, with less than 2 negative interactions between adult and child? *Describe what happened:*	YES NO
Did the session last no more than 5 **minutes**?	YES NO
Did the adult make the **correct number of comments (zero)**?	YES NO
Did the adult include the **correct number of instructional trials (zero)**?	YES NO

PROCEDURAL FIDELITY – Descriptive comments intervention session (DC)- continued	
Child ID #:_____Researcher initials:_____ Date: _____ Session #: ___	
SET UP / IMPLEMENTATION	(circle one)
Were the adult and child in a **one on one** setting?	YES NO
Did the adult provide **child choice** of activity?	YES NO
Did the session have no more than three **disruptive interruptions / interference** from peers?	YES NO
Was the session conducted in a **positive manner**, with less than 2 negative interactions between adult and child?	YES NO
Did the session last between **8-10 minutes**?	YES NO
Did the adult make the **correct number of commentsfor target letter 1**? (10-12)	YES NO
Did the adult make the **correct number of commentsfor target letter 2**? (10-12)	YES NO
Did the adult include the **correct number of instructional trials for target letter 1** (zero)?	YES NO
Did the adult include the **correct number of instructional trials for target letter 2** (zero)?	YES NO

Was there anything **unusual** about the session? (child was returning from being sick for several days, had just returned from a field trip, etc)

_____

Comments (list in order occurred with time heard on audio)

	target letter	saidletter?	included descriptor?	total		target letter	saidletter?	included descriptor?	total
1				/2	15				/2
2				/2	16				/2
3				/2	17				/2
4				/2	18				/2
5				/2	19				/2
6				/2	20				/2
7				/2	21				/2
8				/2	22				/2
9				/2	23				/2
10				/2	24				/2
11				/2	25				/2
12				/2	26				/2
13				/2	27				/2
14				/2	28				/2

TOTAL PF:_____ / _____

PROCEDURAL FIDELITY – Descriptive comments intervention session	
	DESCRIPTIVE COMMENTS MADE
TARGET LETTER (set 1)	*10-12 comments per target letter must be made - cross off when completed*
	12 11 10 9 8 7 6 5 4 3 2 1
	12 11 10 9 8 7 6 5 4 3 2 1
(CONTINUED OTHER SIDE)	

CHILD TARGET PRODUCTIONS

Target letter	time	code (SP or IM)

PROCEDURAL FIDELITY – Descriptive comments plus trials intervention (DC + T)- continued	
Child ID #:_____Researcher initials:_____ Date: _____ Session #: _____	
SET UP / IMPLEMENTATION	(circle one)
Were the adult and child in a **one on one** setting?	YES NO
Did the adult provide **child choice** of activity?	YES NO
Did the session have no more than three **disruptive interruptions / interference** from peers?	YES NO
Was the session conducted in a **positive manner**, with less than 2 negative interactions between adult and child?	YES NO
Did the session last between **8-10 minutes**?	YES NO
Did the adult make the **correct number of comments for target letter 1**? (2)	YES NO
Did the adult make the **correct number of comments for target letter 2**? (2)	YES NO
Did the adult include the **correct number of instructional trials for target letter 1**? (4)	YES NO
Did the adult include the **correct number of instructional trials for target letter 2**? (4)	YES NO

Was there anything **unusual** about the session? (child was returning from being sick for several days, had just returned from a field trip, etc)

_____

Descriptive Comments

	Target letter	said letter?	included descriptor?	total
1				/ 2
2				/ 2
3				/ 2
4				/ 2
5				/ 2
6				/ 2
			TOTAL PF	/

PROCEDURAL FIDELITY – Descriptive comments plus trials intervention (DC + T)

COMMENTS MADE

TARGET LETTER (set 2)	*check box when completed*		*tally*	
	1 given before trials begin	1 given after trials	number made when prompting	TOTAL

CHILD TARGET PRODUCTIONS

Target letter	time	code (SP or IM)

PROCEDURAL FIDELITY -CTD Trials (0 sec)

Instructions: Mark ✓ if behavior occurred, × if it did not, and — if behavior was not applicable.

	Target letter	Provided task direction "What letter is this?"	Immediately gave controlling prompt "Say B"	Gives descriptive praise with letter ("You got it – that's a B!") after correct response	Gives feedback with letter and moves on "No, it's a B") after incorrect or no response (after 3 sec.)	Total present # of✓ / # of✓and ×
Trial 1						/
Trial 2						/
Trial 3						/
Trial 4						/
Trial 5						/
Trial 6						/
Trial 7						/
Trial 8						/
Trial 9						/
Trial 10						/
					TOTAL PF	/

PROCEDURAL FIDELTY CTD Trials 3-sec

Instructions: Mark + if behavior occurred, 0 if it did not, and — if behavior was not applicable.

	Target letter	Provided task direction "What letter is this?"	Waited correct amt of time (3 sec)	If correct, adult gives descriptive praise with letter ("You got it – that's a B!")	If incorrect or no response after 3sec., adult delivers controlling prompt ("No. Say B")	After controlling prompt and correct response, adult gives descriptive praise with letter	After controlling prompt and incorrect or NR, adult gives feedback, states letter, and moves on ("No, It's a B")	Total present # of + /# of + and 0
Trial 1								/
Trial 2								/
Trial 3								/
Trial 4								/
Trial 5								/
Trial 6								/
Trial 7								/
Trial 8								/
Trial 9								/
Trial 10								/
							TOTAL PF	/

PROCEDURAL FIDELITY – Generalization across materials probe	
SET UP / IMPLEMENTATION	(circle one)
Were the adult and child in a **one on one** setting?	YES NO
Did the session have no more than three **disruptive interruptions / interference** from peers? *Describe:*	YES NO
Was the session conducted in a **positive manner**, with less than 2 negative interactions between adult and child? *Describe:*	YES NO
Did the session last between **8-10 minutes**?	
Did the adult make the **correct number of comments (zero)**?	YES NO
Did the adult include the **correct number of instructional trials (zero)**?	YES NO
Did the adult include **3 probe trials per letter** (18 total)?	YES NO
Did the adult **follow the procedure** correctly?	YES NO

Was there anything **unusual** about the session? (child was returning from being sick for several days, had just returned from a field trip, etc)

_____

_____

PROCEDURAL FIDELITY – Generalization across adults	
SET UP / IMPLEMENTATION	(circle one)
Were the adult and child in a **one on one** setting?	YES NO
Did the adult provide **child choice** of activity?	YES NO
Did the session have no more than three **disruptive interruptions / interference** from peers? *Describe:*	YES NO
Was the session conducted in a **positive manner**, with less than 2 negative interactions between adult and child? *Describe:*	YES NO
Did the session last between **8-10 minutes**?	YES NO
Did the adult make the **correct number of comments (zero)**?	YES NO
Did the adult include the **correct number of instructional trials (zero)**?	YES NO
Did the adult include **3 probe trials per letter**?	YES NO

Was there anything **unusual** about the session? (child was returning from being sick for several days, had just returned from a field trip, etc)

_____

_____

REFERENCES

Alig-Cybriwsky, C., Wolery, M., & Gast, D. L. (1990). Use of a constant time delay procedure in teaching preschoolers in a group format. *Journal of Early Intervention, 14,* 99-116.

Bambara, L. M., Warren, S. F., & Komisar, S. (1988). The individualized curriculum sequencing model: Effects on skill acquisition and generalization. *The Journal for the Association for Persons with Severe Handicaps, 13,* 8-19.

Barnett, S., Epstein, D. J., Friedman, A. H., Boyd, J. S., & Hustedt, J. T. (2008). *The state of Preschool: 2008 state preschool yearbook*. New Brunswick, NJ: The National Institute for Early Education Research.

Barton, E., & Wolery, M. (2007) Evaluation of e-mail feedback on the verbal behaviors of pre-service teachers. *Journal of Early Intervention, 30,* 55-72.

Bricker, D. D. (1989). *Early intervention for at-risk and handicapped infants, toddlers, and preschool children.* Palo Alto, CA: VORT Corporation.

Chiara, L., Schuster, J.W., Bell, J. K., & Wolery, M. (1995). Small group massed-trial and individually-distributed-trial instruction with preschoolers. *Journal of Early Intervention,19,* 203-217.

Clements, D. H., Sarama, J., & DiBiase, A. M. (2004). *Engaging young children in mathematics: Standards for early childhood mathematics education.* Mahwah, NJ: Lawrence Erlbaum.

Cole, K. N. & Dale, P. S. (1986). Direct language instruction and interactive language instruction with language delayed preschool children: A comparison study. *Journal of Speech and Hearing Research, 29,* 206-217.

Copple, C., & Bredekamp, S. (Eds.). (2009). *Developmentally appropriate practice in early childhood programs.* Washington, DC: National Association for the Education of Young Children.

Daugherty, S., Grisham-Brown, J., & Hemmeter, M. L. (2001). The effects of embedded skill instruction on the acquisition of target and nontarget skills in preschoolers with developmental delays. *Topics in Early Childhood Special Education, 21,* 213-221.

DeThorne, L. S., Johnson, C. J., Walder, L, & Mahurin-Smith, J. (2009). When "Simon Says" doesn't work: Alternatives to imitation for facilitating early speech development. *American Journal of Speech-Language Pathology, 18*, 133-145.

Doyle, P. M., Wolery, M., Gast, D. L., & Ault, M. J. (1990). Comparison of constant time delay

and the system of least prompts in teaching preschoolers with developmental delays. *Research in Developmental Disabilities, 11*, 1-22.

Dunn, L.M., & Dunn, L.M. (1997). *Peabody Picture Vocabulary Test.* Circle Pines, MN: American Guidance Service.

Ellis Weismer, S., Murray-Branch, J., & Miller, J. F. (1993). Comparison of two methods for promoting vocabulary in late talkers. *Journal of Speech and Hearing Research, 36,* 1037-1050.

Ellis Weismer, S., & Roberstson, S. (2007). Focused stimulation approach to language intervention. In R.J., McCauley & M. E. Fey (Eds.) *Treatment of language disorders in children* (pp. 175-202). Baltimore: Paul H. Brooks.

Englemann, S. & Osborn, J. (1976). *DISTAR: Language I: An Instructional System.* Chicago: Science Research Associates.

Fenson, L., Dale, P.S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., et al. (1993). Technical Manual for the MacArthur Communicative Development Inventories. San Diego, CA: Singular.

Fey, M. (1986). *Language intervention with young children.* San Diego: College-Hill Press.

Fox, L., & Hanline, M.F. (1993). A preliminary evaluation of learning within developmentally appropriate early childhood settings. *Topics in Early Childhood Special Education, 13,* 308-327.

Gardner, E. (1979). *Expressive One Word Picture Vocabulary Test.* Novado, CA: Academic Therapy.

Girolametto, L., Pearce, P. S., & Weitzman, E. (1996). Interactive focused stimulation for toddlers with expressive vocabulary delays. *Journal of Speech and Hearing Research, 39,* 1274-1283.

Girolametto, L., Weitzman, E., Clements-Baartman, J. (1998). Vocabulary intervention for children with Down Syndrome: Parent training using focused stimulation. *Infant-Toddler Intervention, 8,* 109-125.

Grisham-Brown, J., Schuster, J. W., Hemmeter, M. L., & Collins, B. C. (2000). Using an embedding strategy to teach preschoolers with significant disabilities. *Journal of Behavioral Education, 10,* 139-162.

Grisham-Brown, J. (2008). Best practices in implementing standards in early childhood education. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 1025–1042). Washington, DC: National Association of School Psychologists.

Grisham-Brown, J., Pretti-Frontczak, K., Hawkins, S. R., & Winchell, B. N. (2009). Addressing early learning standards for all children within blended preschool classrooms. *Topics in Early Childhood Special Education, 29,* 131-142.

Guess, D., Horner, R. D., Utley, B., Holvoet, J., Maxon, D., Tucker, D., & Warren, S. (1978). A functional curriculum sequencing model for teaching the severely handicapped. *AAESPH Review, 3,* 202-215.

Hancock, T. B., & Kaiser, A. P. (2005). Enhanced Milieu Teaching. In R. McCauley & M. Fey (Eds.) *Treatment of language disorders in children* (pp. 203-236). Baltimore: Paul Brookes.

Harris, M. D. & Reichle, J. (2004). The impact of aided stimulation on symbol comprehension and production in children with moderate cognitive disabilities. *American Journal of Speech Language Pathology, 13*, 155-167.

Heal, N. A., Hanley, G. P., & Layer, S. A. (2009). An evaluation of the relative efficacy of and children's preference for teaching strategies that differ in amount of teacher directedness. *Journal of Applied Behavior Analysis, 42,* 123-143.

Holcombe, A., Wolery, M., & Snyder, E. (1994). Effects of two levels of procedural fidelity with constant time delay on children's learning. *Journal of Behavioral Education, 4,* 49-73.

Holvoet, J., Guess, D., Mulligan, M., & Brown, F. (1980). The individualized curriculum sequencing model (II): A training strategy for severely handicapped students. *Journal of the Association for the Severely Handicapped, 5,* 337-351.

Horn, E., Lieber, K., Li, S., Sandall, S., & Schwartz, I. (2000). Supporting young children's IEP goals in inclusive settings through embedded learning opportunities. *Topics in Early Childhood Special Education, 20,* 208-223.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S. L., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practices in special education. *Exceptional Children, 71*, 165-179.

Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbain, O. (2008) Ready to learn? Children's pre-academic achievement in pre-Kindergarten programs. *Early Childhood Research Quarterly, 23*, 27-50.

Individuals with Disabilities Education Improvement Act of 2004, Pub. L. No. 108-446, U.S.C. 20, 1400 et seq.

Kaiser, A. P., Ostrosky, M. M., & Alpert, C. L. (1993). Training teachers to use environmental arrangement and milieu teaching. *Journal of the Association for People with Severe Handicaps, 18*, 188-199.

Kaiser, A. P. & Trent, J. A. (2007). Communication intervention for young children with disabilities: Naturalistic approaches to promoting development. In S. L. Odom, R. H. Horner, M. E. Snell, & J. Blacher (Eds.), *Handbook of Developmental Disabilities*, (pp. 224-245). New York, NY: Guilford Press.

Kazdin, A. E. (1982). *Single-case research designs: Methods for clinical and applied settings.* New York: Oxford University Press.

Kinder, K. A. (2009).*Using embedding instruction to teach pre-academic skills: Comparing interventions and outcomes.* Unpublished manuscript, Vanderbilt University, Nashville, TN.

Kouri, T. A. (2005). Lexical training through modeling and elicitation procedure with late talkers who have specific language impairment and developmental delays. *Journal of Speech Language and Hearing Research,48*, 157-171.

Leonard, L.B. (1981). Facilitating linguistic skills in children with specific language impairment. *Applied Psycholinguistics,2,* 89-188.

Losardo, A., & Bricker, D. (1994). Activity-based intervention and direct instruction: A comparison study. *American Journal on Mental Retardation, 98*, 744-765.

MacDonald, J. D. (1989). *Becoming partners with children: from play to conversation. A developmental guide for professionals and parents.* Chicago: Riverside.

Mahoney, G., & Powell, A. (1988). Modifying parent-child interaction: Enhancing the development of handicapped children. *The Journal of Special Education, 22,* 82-96.

McBride, B. J., & Schwartz, I. S. (2003). Effects of teaching early interventionists to use discrete trials during ongoing classroom activities. *Topics in Early Childhood Special Education, 23,* 5-17.

Mullen, E. M. (1995). *Mullen Scales of Early Learning* (AGS ed.). Circle Pines, MN: American Guidance Service, Inc.

Nolet, V. & McLaughlin, M. J. (2005). *Accessing the general curriculum: Including students with disabilities in standards based reform*. Thousand Oaks, CA: Corwin Press.

Peck, C. A., Killen, C. C., & Baumgart, D. (1989). Increasing implementation of special education instruction in mainstream preschools: Direct and generalized effects of nondirective consultation. *Journal of Applied Behavior Analysis, 22,*197-210.

Pemberton, E. F., & Watkins, R. V. (1987). Language facilitation through stories: Recasting and modeling. *First Language, 7,* 79-89.

Pretti-Frontczak, K. & Bricker, D. (2004). *An activity-based approach to early intervention.* Baltimore: Brookes.

Reid, D. K., Hresko, W. P., Hammill, D. D. (2001). *Test of early reading ability.* Austin, TX: Pro-Ed, inc.

Riches, N. G., Tomasello, M, & Conti-Ramsden, G. (2005). Verb learning in children with SLI: Frequency and spacing effects. *Journal of Speech Language and Hearing Research, 48,* 1397-1411.

Rice, M. L. (1993). "Don't talk to him, he's weird": A social consequences account of language and social interactions. In S.F. Warren & J. Reichle (Series Eds.) & A.P. Kaiser & D. Gray (Vol. Eds.), *Communication and language intervention series: Vol. 2. Enhancing children's communication: Research foundation for early language intervention* (pp. 139-158). Baltimore: Brookes.

Roberston, S. B. & Ellis Weismer, S. (1999). Effects of treatment on linguistic and social skills in toddlers with delayed language development. *Journal of Speech, Language, and Hearing Research, 42,* 1234-1248.

Rogers-Warren, A. K., & Warren, S. F. (1980). Mands for verbalization: Facilitating the generalization of newly trained language in children. *Behavior Modification, 4,* 230-245.

Rule, S., Losardo, A., Dinnebiel, L., Kaiser, A., & Rowland, C. (1998). Translating research on naturalistic instruction into practice. *Journal of Early Intervention, 21,* 283-293.

Salmon, C. M., Rowan, L. E., & Mitchell, P. R. (1998).Facilitating prelinguistic communication: Impact of adult prompting. *Infant-Toddler Intervention, 8,* 11-27.

Sandall, S. R. & Schwartz, I. S. (2002). Building blocks for teaching preschoolers with special needs. Baltimore: Brookes.

Schepis, M. M., Reid, D. H., Ownbey, J., & Parsons, M. B. (2001). Training support staff to embed teaching within natural routines of young children with disabilities in an inclusive preschool. *Journal of Applied Behavior Analysis, 34,* 313-327.

Scott-Little, C., Kagan, S. L., & Frelow, V. S. (2003). Creating the conditions for success with early learning standards: Results from a national study of state-level standards for children's learning prior to kindergarten*. Early Childhood Research & Practice, 5.* Retrieved from http://ecrp.uiuc.edu/v5n2/little.html

Scott-Little, C., Kagan, S. L., & Frelow, V. S. (2006). Conceptualization of readiness and the content of early learning standards: The intersection of policy and research. *Early Childhood Research Quarterly, 21,* 153-173.

Sewell, T. J., Collins, B. C., Hemmeter, M. L., & Schuster, J. W. (1998). Using simultaneous prompting within an activity-based format to teach dressing skills to preschoolers with developmental delays. *Journal of Early Intervention, 21,* 132-145.

Sindelar, P. T., Rosenberg, M. S., & Wilson, R. J. (1985). An adapted alternating treatments design for instructional research. *Education & Treatment of Children, 8*, 67-76.

Smith, M. W., & Dickinson, D. K. (2002). *User's guide to the Early Language and Literacy Classroom Observation (ELLCO)*. Baltimore: Brookes Publishing.

Snell, M. E. (2007). Advances in instruction. In S. L. Odom, R. H. Horner, M. E. Snell, and Martha E., & J. Blacher,(Eds), *Handbook of developmental disabilities* (pp 249-268). New York: Guilford Press.

Sorensen, P., & Fey, M. (1992). Informativeness as a clinical principle: What's really new? *Language, Speech, and Hearing Services in Schools, 23,* 320-328.

Snyder, P., Artman, K., Hemmeter, M. L., Kinder, K., & Pasia, C. (2010). Characterizing key features of the early childhood professional development literature. Manuscript in preparation.

Tekin, E., & Kircaali-Iftar, G. (2002). Comparison of the effectiveness and efficiency of two response prompting procedures delivered by sibling tutors. *Education and Training in Mental Retardation and Developmental Disabilities, 37*, 283-99.

Warren, S. F., & Kaiser, A. P. (1986). Incidental language teaching: A critical review. *Journal of Speech & Hearing Disorders, 51,* 291-299.

Werts, M. G., Wolery, M., Holcombe-Ligon, A., Vassilaros, M. A., & Billings, S. S. (1992). Efficacy of transition-based teaching with instructive feedback. *Education & Treatment of Children, 15,* 320-334.

Wilcox, M. J., Kouri, T. A., & Caswell, S. B. (1991). Early language intervention: A comparison of classroom and individual treatment. *American Journal of Speech and Language Research*, 49-62.

Wolery, M. (2005). DEC Recommended practices: Child focused practices. In S. Sandall, M. L. Hemmeter, B. J. Smith, & M. E. McLean (Eds.), *DEC recommended practices: A comprehensive guide for practical application* (pp. 71–106). Longmont, CO: Sopris West.

Wolery, M., Ault, M. J., & Doyle, P. M. (1992). *Teaching students with moderate to severe disabilities: Use of response prompting strategies.* White Plains, NY: Longman.

Wolery, M., Anthony, L., & Heckathorn, J. (1998).Transition-based teaching: Effects on transitions, teachers' behavior, and children's learning. *Journal of Early Intervention, 21,* 117-131.

Wolery, M., Anthony, L., Caldwell, N. K., Snyder, E. D., & Morgante, J. D. (2002). Embedding and distributing constant time delay in circle time and transitions. *Topics in Early Childhood Special Education, 22,* 14-25.

Wolery, M., Doyle, P. M., Gast, D. L., Ault, M. J., & Simpson, S. L. (1993). Comparison of progressive time delay and transition-based teaching with preschoolers who have developmental delays. *Journal of Early Intervention, 17,* 160-176.

Wolery, M. & Dunlap, G. (2001). Reporting on studies using single-subject experimental methods. *Journal of Early Intervention, 24*, 85-89.

Wolery, M., Gast, D. L., & Hammond, D. (2010). Comparative intervention designs. In David L. Gast (Ed.), *Single subject research methodology in behavioral sciences* (pp. 329-381). New York: Routledge.

Wolery, M., Holcombe, A., Cybriwsky, C., Doyle, P. M., Schuster, J. W., & Ault, M. J. (1992). Constant time delay with discrete responses: A review of effectiveness and demographic, procedural, and methodological parameters. *Research in Developmental Disabilities,13,* 239-266.

Yoder, P. J., Kaiser, A. P., & Alpert, C. (1991). An exploratory study of the interaction between language teaching methods and child characteristics. *Journal of Speech and Hearing Research, 34,* 155-167.

Yoder, P. J, Kaiser, A. P., Goldstein, H., Alpert, C., Mousetis, L., Kaczmarek, L., & Fischer, R. (1995). An exploratory comparison of milieu teaching and responsive interaction in classroom applications. *Journal of Early Intervention, 19,* 218-242.

Yoder, P. J. & Warren, S. F. (2001). Relative treatment effects of two prelinguistic communication interventions on language development in toddlers with developmental delays vary by maternal characteristics. *Journal of Speech Language Hearing Research, 44,* 224-237.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2002). *Preschool language scale, fourth edition.* San Antonio, TX: Harcourt Assessment.

Ziolkowski, R. A., & Goldstein, H. (2008). Effects of an embedded phonological awareness

intervention with repeated book reading on preschool children with language delays. *Journal of Early Intervention, 31,* 67-90.