Research in Autism Spectrum Disorders 5 (2011) 267–276

Contents lists available at ScienceDirect

Research in Autism Spectrum Disorders

Journal homepage: http://ees.elsevier.com/RASD/default.asp

Addressing working memory in children with autism through behavioral intervention☆

Lisa Baltruschat [a], Marcus Hasselhorn [b], Jonathan Tarbox [c,*], Dennis R. Dixon [c], Adel C. Najdowski [c], Ryan D. Mullins [c], Evelyn R. Gould [c]

[a] Center of Research on Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt/M and the Center for Autism and Related Disorders, Germany
[b] German Institute for International Educational Research, Center of Research on Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt/M, Germany
[c] Center for Autism and Related Disorders, Tarzana, CA 91356, United States

ARTICLE INFO

Article history:
Received 3 March 2010
Accepted 12 April 2010

Keywords:
Autism
Executive functions
Working memory
Positive reinforcement
Central-executive
Applied behavior analysis

ABSTRACT

Children with autism often struggle with executive function (EF) deficits, particularly with regard to working memory (WM). Despite the documented deficits in these areas, very little controlled research has evaluated treatments for remediation of EF or WM deficits in children with autism. This study examined the use of positive reinforcement for improving performance on Counting Span tasks which are said to measure the central-executive component of working memory. Large improvements in performance were obtained for all participants, as was maintenance and generalization to untrained stimuli and untrained responses. Results suggest that basic behavioral intervention procedures may be successful in improving performance on complex behaviors such as those labeled as "working memory."

Autism is a pervasive developmental disorder characterized by impairments in communication, social interaction, and the presence of restricted, repetitive and stereotypical patterns of behavior and interests (American Psychiatric Association; APA, 1994). Children displaying an autism spectrum disorder (ASD) may have delays in many areas such as cognition, social skills, language, play, adaptive, and motor skills (Tervo, 2003). Among the many potential areas of deficit in ASD is executive function (Hughes, 2001; Ozonoff, 1995a, 1995b; Ozonoff & Jensen, 1999). Furthermore, deficits in executive function (EF) may be related to skill deficits in other areas such as inflexibility, perseveration and inappropriate responding to social situations (Bennetto, Pennington, & Rogers, 1996).

Executive function is a concept used to describe a loosely defined collection of brain processes which are said to be responsible for planning, goal persistency, cognitive flexibility, abstract thinking, rule acquisition, initiating appropriate actions and inhibiting inappropriate actions, and selecting relevant sensory information (Hill, 2004). However, the term EF lacks a precise operational or conceptual definition between and within the various scientific disciplines (Ruble & Scott, 2002). Indeed, Daffner and Searl (2008) write that "There is no commonly accepted definition or model of EF, which makes a discussion at the construct level a challenge." A common psychologically weighted definition describes EF as "the ability to maintain an appropriate problem solving set for attainment of a future goal" (Ozonoff, Pennington, & Rogers, 1991).

☆ This study was conducted in partial fulfillment of Lisa Baltruschat's doctoral degree in psychology. The authors would like to thank Andrew Kaplan, Sarah Tuberman, and Pearly Cheng for their help with data collection.

* Corresponding author at: Center for Autism and Related Disorders, 19019 Ventura Blvd, 3rd Floor, Tarzana, CA 91356, United States. Tel.: +1 818 345 2345.
E-mail address: j.tarbox@centerforautism.com (J. Tarbox).

1750-9467/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.rasd.2010.04.008

Neuroscientists often treat EF as an umbrella term, comprising a very broad network of cognitive and behavioral skills and processes required to plan, initiate and follow through activities. For example, one definition describes EF as ‘‘a complex set of cerebral processes that operate in non-routine situations and exert top–down, volitional control over cognition and behavior’’ (Daffner & Searl, 2008). Furthermore, EF is often assumed to be housed in particular cerebral structures, such as the prefrontal lobes of the brain (Daffner & Searl, 2008; Dawson & Guare, 2004).

One area of EF is which is frequently discussed in ASDs is working memory (WM). WM is considered a component of EF due to its function in the organizational aspects of memory and the role it plays in goal-directed behavior (Pennington & Ozonoff, 1996). The term WM describes the ability to store (‘‘keep online’’) information and process the information at the same time. Most researchers agree that WM consists of several specialized components. However, there is little agreement on the exact nature and composition of the components (Alloway, Gathercole, & Pickering, 2006). Baddeley and Hitch proposed a detailed model of WM in 1974, consisting of separate memory systems which are responsible for the temporary storage of visuospatial and phonological representations (Baddeley, 2000). According to Baddeley's model, WM consists of three components: the visuospatial sketchpad, the phonological loop, and the central-executive, all of which work together to process incoming information. The visuospatial sketchpad governs the processing of visual information, the phonological loop auditory information, and the central-executive which acts as a supervisory system and controls the flow of information from the other two systems. An additional component was added in 2000, known as the ‘episodic buffer.’ It links the visual, verbal and auditory information within the correct chronological sequence.

Research has found evidence of working memory deficits in individuals with ASD across a wide range of chronological and mental ages (Geurts, Verte, Oosterlaan, Roeyers, & Sergeant, 2004; Ozonoff, 1997; Verté, Geurts, Roeyers, Oosterlaan, & Sergeant, 2006; see Hill, 2004 for a recent review). A significant amount of research has been done on deficits in the central-executive in children with ASDs, in particular (Ozonoff et al., 1991; Prior & Hoffman, 1990; Rumsey & Hamburger, 1988). As described above, the central-executive is said to coordinate the other two systems (visouspatial sketchpad and phonological loop), and helps bind information from a number of sources into coherent episodes. Shifting between tasks, retrieving new strategies, inhibiting inappropriate reactions, and strengthening selective attention are all said to be tasks of the central-executive.

Despite the documented deficits in EF and WM in ASDs, relatively little research has been published on procedures for remediating these deficits. A small number of studies have evaluated approaches to improving working memory, often focusing on children with ADHD, FASD or Down syndrome. Tamm et al. (2009) examined whether training selective, alternating, and divided attention would produce EF gains in children with ADHD. They found that the children's performance increased on neuropsychological tests in EF areas including WM. Farb and Throne (1978) found that a rehearsal training program effectively improved the mnemonic performance of a child with Down syndrome. This finding was later replicated with typically developing peers. For example, Turley-Ames and Whitfield (2003) found that WM span scores increased as a result of using a rehearsal strategy. However, little or no research has been published on intervention for WM deficits in individuals with ASD.

Little attention has been paid in the behavior analytic literature to the topics of EF and WM (however, see Hayes, Gifford, & Ruckstuhl, 1996, for a conceptual treatment). A small number of basic studies conducted with typically developing adults demonstrated that positive reinforcement can affect performance on short-term memory tests (Cuvo, 1974; Loftus, 1972) but very little research has been published since then and virtually no research has been done on clinical applications of behavioral procedures. A common misconception of the behavioral orientation is that it rejects the events which are commonly labeled as memory and cognition from scientific inquiry, but this is not the case. From a behavioral perspective, anything anyone does in interacting with their environment is to be considered behavior and behavior is assumed to be amenable to change via basic processes of learning and motivation. Thus, no distinction is made between ‘‘cognitive’’ activities which occur internally and behaviors which occur externally. Rather, the distinction which is generally made is between ‘‘private’’ or covert behavior and ‘‘public’’ or overt behavior. The only defining difference between public and private events is that public are amenable to observation by more than one person simultaneously, whereas private events can only be observed by the person engaging in them, if at all (Skinner, 1974). Despite the difference in observability between public and private events, it is assumed that the two types of behavior are not different in kind. They are presumably acquired in the same manner (via behavioral principles of learning) and they obey the same laws of interaction with the environment (e.g., reinforcement, extinction, stimulus control, stimulus generalization, etc.). This philosophical position is referred to as Radical Behaviorism (Skinner, 1974) and it serves as the basic philosophical foundation for contemporary behavior analytic science.

Aside from whether one agrees with the philosophical and theoretical premises of Radical Behaviorism, taking such an approach to working memory is inherently practical because it provides recommendations for treatment that are concrete and easily testable. For example, if working memory performance involves behavior, then one should be able to improve working memory performance through the most foundational of behavioral processes: positive reinforcement. Furthermore, if mediating behaviors, such as rehearsal, play a critical part in working memory, then basic behavioral treatment procedures such as prompting and reinforcement should be useful for teaching these skills in people who do not have them. Given that behavioral intervention procedures are based on established principles of learning (Catania, 1998) and have been well-validated as treatment procedures in many areas, including autism (Granpeesheh, Tarbox, & Dixon, 2009), it should be relatively easy to identify which procedures are likely to be effective. Further, there already exists an extensive research literature that provides the basic methodology for implementing and testing them.

Despite the potential utility of developing a behavioral treatment for EF and WM deficits in individuals with ASD, virtually no research has been published in this area. Therefore, the purpose of this study was to evaluate a basic behavioral procedure for improving some aspect of EF, and more specifically WM, in children with ASD. As discussed previously, the central-executive is a component of WM which has been shown to impaired in individuals with ASD and was therefore selected as a target of intervention in this study. The "Counting Span" (CS) task is a commonly used procedure for assessing the central-executive of children (Andersson, 2008; Bull, Johnson, & Roy, 1999; Fournet, Morand, Roulin, Naegele, & Pellat, 1996). The task involves presenting a series of cards with quantities of visual stimuli (e.g., colored dots) and the participant is told to count the quantity each time another card is presented, and then eventually recall all the quantities in the correct order. The task is said to measure a participant's ability to simultaneously store and process information and it was developed to prevent the participant from rehearsing the stimuli in order to memorize it, that is, the frequent demand to count new stimuli is presumed to interfere with rehearsal. In this study, performance on the CS task was directly reinforced in children with ASD. Further, maintenance was assessed after positive reinforcement was terminated, as was generalization to novel stimuli and responses, thereby allowing for an assessment of whether positive reinforcement resulted in improvements in participants' generalized ability on CS tasks, rather than memorization of particular behaviors.

1. Method

1.1. Participants and setting

The criteria for recruitment for participation in the study included: (1) minimum of 6 years of age, (2) diagnosis on the autism spectrum, (3) have relatively well-developed language repertoires (ability to understand rules, follow complex instructions, etc.), (4) and that the child's family and therapy team expressed that improvement in the general area of executive function – and working memory in particular – was a clinical priority. Three boys with autistic disorder participated in the study. All were current clients of a large-scale, community-based provider of home-based behavioral intervention services. All were receiving comprehensive behavioral intervention services that addressed all skill areas in which the boys demonstrated deficits (e.g., academics, social skills, language, play, adaptive, and motor skills). All study procedures were approved by an institutional review board and all participants' parents gave their informed consent in writing before the study began.

Adam was 9 years old, had received behavioral intervention services for approximately 6 years, and was receiving between 9 and 16 h per week of behavioral intervention at the time of the study. Adam was in good general health, with no history of seizures, accidents or hospitalizations. At the time of the study, Adam was taking 25 mg of Zoloft and 1 mg of Tenex per day and continued to do so throughout the study.

Joe was 7 years old and had been receiving behavioral intervention for 5 years. At the time of the study, Joe was receiving 35 h per week of behavioral intervention, consisting of 20 h per week of school shadowing and 15 h per week of in-home therapy. Joe was in good general health, with no history of seizures, accidents or hospitalization. At the time of the study, Joe was on a gluten and casein free diet and remained on it throughout the study.

Dave was 11 years old, had received behavioral intervention for 15 months prior to the start of the study, and was receiving 10 h per week of behavioral intervention throughout the time of the study. Dave had experienced infrequent grand mal seizures (i.e., 3 total) during the year prior to the study and was therefore receiving a gradually increasing dose of Lamictal during the course of the study. The final dose achieved and maintained during the latter part of the study was 150 mg per day.

All sessions during *baseline*, *positive reinforcement*, *maintenance*, and *generalization* conditions were conducted in the participants' home environments, where their behavioral intervention sessions normally took place (usually their bedroom or spare room). The work area was arranged in an age appropriate manner for the child and included a desk, chairs, and a variety of intervention materials. Sessions were conducted two-to-four times per day, two-to-three days per week.

1.2. Experimental design

Since individual differences among the participants were expected, a single-subject design was used in order to evaluate and detect the effects of the intervention at the level of the individual child. Since it was hoped that the treatment effects would maintain after treatment was withdrawn (i.e., the effects would not reverse), a multiple baseline across participants design was implemented.

1.3. Response measurement and interobserver agreement

Data were collected on the accuracy of responding for each trial during *baseline*, *positive reinforcement*, *maintenance*, and *generalization* phases. A correct response was defined as the participant accurately stating the quantities presented, in the order in which they were presented. Data were graphed and analyzed as percentage correct.

A second trained observer independently collected data during 85%, 86%, and 84% of sessions for Adam, Joe, and Dave, respectively. Interobserver agreement (IOA) was calculated by dividing the number of trials for which both observers scored exactly the same data by the total number of trials for which two observers scored data, and the resulting decimal was multiplied by 100, thereby converting it into a percentage. Mean IOA was 91% (range = 80–100%), 86% (range = 80–100%), and 99% (range = 80–100%) for Adam, Joe, and Dave, respectively.

1.4. Arbeitsgedaechtnis Testbatterie (AGTB)

All participants completed the Arbeitsgedaechtnis Testbatterie (AGTB), translated as "Working Memory Test Battery," at the outset of the study, as well as at the very end. The AGTB is a computerized evaluation, using a touch-sensitive monitor, which assesses the central-executive WM of children (Hasselhorn et al., in press). The AGTB is comprised of six subtests, including: (1) *Complex Span*, (2) *Color Span*, (3) *Digit Span Backwards*, (4) *Stroop-Like*, (5) *Go/No Go*, and (6) *Counting Span*. The AGTB has been translated into English for use with English speaking participants.

1.4.1. Subtests

During the *Complex Span* subtest, a sequence of visual stimuli was presented. For each stimulus, the participant is asked to emit a classification response according to the function of the object (i.e., "can you eat it?"). At the end of the sequence of the stimuli, the participant was asked to state the names of the stimuli in the order they had been presented. During the *Color Span* subtest, the participants were asked to memorize visually presented colors and recall them backwards by selecting the correct ones out of a visually presented ring of circles, each of which was a different color. During the *Digit Span Backwards* subtest, the participant had to memorize a series of numbers that are presented auditorally and then recall them backwards. During the *Stroop-Like* subtest, visual and auditory stimuli were presented which were either congruent or incongruent and the participant was asked to respond only to the congruent stimuli, thereby inhibiting their response to the incongruent stimuli. During the *Go/No Go* subtest, the participant was given a rule as to which stimuli he/she should try to detect and then given the opportunity to press a button each time they detect it. During the *Counting Span* subtest, the participant was required to count quantities of circles mixed up with squares in a series of visual arrays (see top image of Fig. 1). The participants were asked to state the quantity each set time an array was presented and were then required to recall each quantity in the correct order in which they were presented.

All AGTB assessment sessions were conducted in a room at the clinic. The room contained two desks, two chairs, and a computer fitted with a touch-screen monitor. The duration of each AGTB session was approximately 1 h. AGTB sessions were conducted one time at the very beginning of the study and one time at the very end.

1.5. Materials

The middle and lower images of Fig. 1 illustrate the materials used during table top sessions (i.e., all sessions aside from the AGTB). Materials were developed in order to measure the same ability as the *Counting Span* subtest on the AGTB but with different stimuli. The table top materials used different shapes which were of different colors than those in the AGTB. Stimuli were presented on 10 cm × 10 cm flashcards. The flashcards contained 17 green shapes on a white background. The green shapes were a mixture of ovals and triangles with each card varying in the ratio of triangles to ovals. The number of ovals ranged from 1 to 9. The numbers 1–9 were first randomized and then split in half in order to use half of them during training (quantities 1, 5, 9, 4, and 3) and half of them during the generalization probes (quantities 2, 5, 6, 7, and 8). Since the numbers 1–9 are an odd number of quantities, the quantity 5 was placed in both the training and generalization sets of stimuli. For the youngest participant (Joe), only numbers 1–4 were used initially. During generalization probes, the materials differed in terms of shape and color; yellow stars and yellow hearts were used in place of green triangles and ovals.

1.6. Procedures

1.6.1. Pre-baseline evaluation

The purpose of the intervention was to increase accuracy on a CS task. However, CS tasks range in difficulty, depending on how many flashcards are presented in the series. Therefore, it was necessary to identify a number of flashcards to present during baseline which produced consistently low levels of accuracy. In order to determine this, trials were presented starting with only two flashcards stimuli and gradually adding on card on each successive trial, until consistently poor accuracy was demonstrated. During each of these trial, the flashcards were presented to the child one at a time. The child was asked to count silently the number of ovals featured on the flashcard and say the counted quantity out loud to the experimenter. The quantity given by the participant did not have to be counted correctly. Because the study was designed to measure working memory abilities, and not the ability to count, an incorrectly counted quantity was accepted. That is, if 3 ovals were displayed on a card and the participant only counted 2 ovals, the final response was still counted as correct, as long as the participant's recall matched their earlier count. If the participant responded correctly to two consecutive trials with two flashcards presented during each, the number of cards presented was increased to three for the following trial. If correct responding again occurred for two consecutive trials, the number of cards was again increased by one. This procedure was continued until the participant made an error on two consecutive trials. The number of flashcards at which errors occurred on two consecutive trials was chosen for inclusion in baseline (4 flashcards for Adam and Joe, 5 flashcards for Dave).

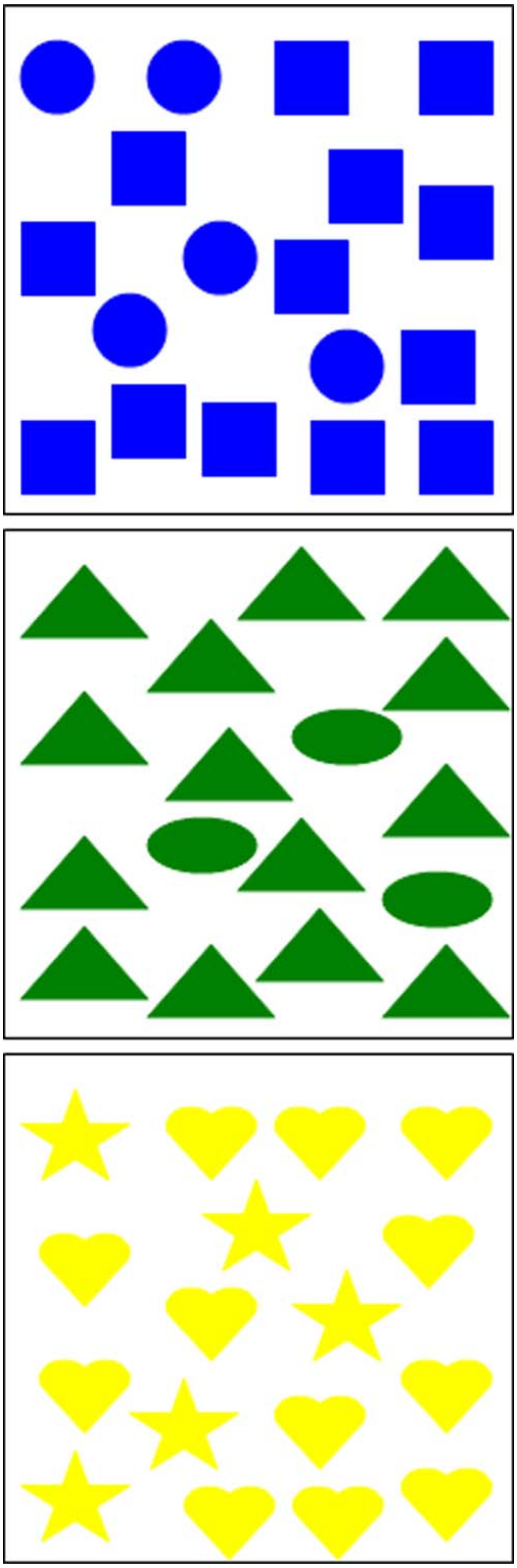

Fig. 1. Sample screenshot during the Counting Span task of the AGTB (top), flashcards used during *baseline* and *positive reinforcement* (middle); and flashcards used during *generalization* (bottom).

1.6.2. Baseline

Before the first baseline trial was conducted, the experimenter stated the instructions to the participant and practiced the task with the child twice. On each baseline trial, flashcards containing quantities of green ovals and triangles were presented. After the child reported the quantity of ovals on the card, the experimenter removed that card and immediately presented the next card, with a different quantity of ovals and triangles. Each time a flashcard was presented to the child the experimenter asked ''How many ovals are there?'' After the final flashcard for a given trial (i.e., at the end of the sequence of flashcards), the experimenter asked ''What do you remember, in the correct order?'' No differential consequences or feedback were provided for correct or incorrect responding, the experimenter simply said ''Okay, let's do the next one,'' and moved on to the next trial.

1.6.3. Positive reinforcement

Prior to each session, the child was asked to select a highly preferred item from an array (e.g., video game, candy, movie, drawing material, or stickers). All items presented were items that participants had restricted access to. That is, participants' parents were asked to not allow participants to access these items outside of the experiment. At the beginning of each session, the procedure was explained to the child. The child was told that for every correct response they would have access to their highly preferred item for 1 min or in the case of a food item; until they had consumed it. If the child did not recall the presented stimuli correctly, they were given corrective feedback (e.g., a neutral ''no, that's not quite right'') and the next trial was conducted. Participants were given 5 min breaks between sessions.

1.6.4. Maintenance

Visual inspection of graphed data was used to determine if the *positive reinforcement* condition resulted in a large and stable increase in correct responding. If it did, participants were then exposed to the maintenance condition. This condition was identical to baseline. Participants were given no indication, whatsoever, of whether they responded correctly or incorrectly.

1.6.5. Generalization

If visual inspection of the data in the *maintenance* phase revealed stable levels of accuracy that were similar to those observed in the *positive reinforcement* condition, participants were exposed to generalization sessions. These sessions were identical to baseline sessions, except that novel stimuli were used that had never been included in the *positive reinforcement* condition. These flashcards included different colors, different shapes, and different quantities to be counted.

2. Results

Fig. 2 depicts the percentage of correct responding during *baseline*, *positive reinforcement*, *maintenance*, and *generalization* sessions, for all participants. The top panel depicts Adam's data. During the baseline phase for Adam, correct responding was consistently low, with baseline/training stimuli (mean = 20%) and generalization stimuli (mean = 20%). When the *positive reinforcement* phase was initiated, Adam's correct responding increased immediately and continued to demonstrate a gradual increase until it remained stable at 100% for 12 consecutive sessions (mean = 83%). The *maintenance* phase was then initiated, wherein reinforcement was discontinued. Adam continued to respond correctly with high accuracy (mean = 90%). He also demonstrated accurate responding with the generalization stimuli, which had never been directly taught or reinforced (mean = 93%).

The middle panel depicts Joe's data. During the initial baseline phase, Joe demonstrated low and variable levels of correct responding to both baseline/training stimuli (mean = 23%) and generalization stimuli (mean = 17%). When positive reinforcement was introduced, Joe's accuracy increased substantially (mean = 74%). Positive reinforcement was then removed during the *maintenance* phase and correct responding remained high to the trained stimuli (mean = 74%) and the untrained generalization stimuli (mean = 80%).

After obtaining maintenance and generalization, the *baseline* phase was reinitiated with stimuli that included quantities 1–9, as opposed to 1–4 which had been implemented with Joe until this point. Joe initially responded favorably during this phase, but his performance rapidly declined (mean = 33%). His correct responses to the generalization stimuli also showed low accuracy (mean = 21%). We therefore initiated the *positive reinforcement* phase and observed an immediate and stable increase in his accuracy (mean = 78%). We then discontinued positive reinforcement in the *maintenance* phase and his accuracy remained high with the trained (mean = 87%) and untrained stimuli (mean = 87%), both of which involved quantities 1–9.

The bottom panel depicts Dave's data, which are similar to those of Adam. During *baseline*, Dave responded with low accuracy to the baseline\training stimuli (mean = 35%) and the generalization stimuli (mean = 33%). He demonstrated an immediate increase during the *positive reinforcement* phase (mean = 92%), maintained correct responding (mean = 78%), as well as demonstrating generalization (mean = 70%), during the *maintenance* phase.

2.1. Secondary analysis: AGTB results

Fig. 3 shows the pre- and post-test scores of all three participants for the *Counting Span* subtest of the AGTB. The *y* axis depicts the mean of the two longest series that were correctly recalled for each participant. This mean was calculated by examining the two trials during the *Counting Span* subtest that had the largest number of quantities which were recalled correctly, and averaging this number for the two trials. Adam's score improved from an average of 3 before treatment to 3.5 cards after treatment. Joe's score changed from an average of 2 before treatment to 3 after treatment. Dave's score increased from 3 to 5.

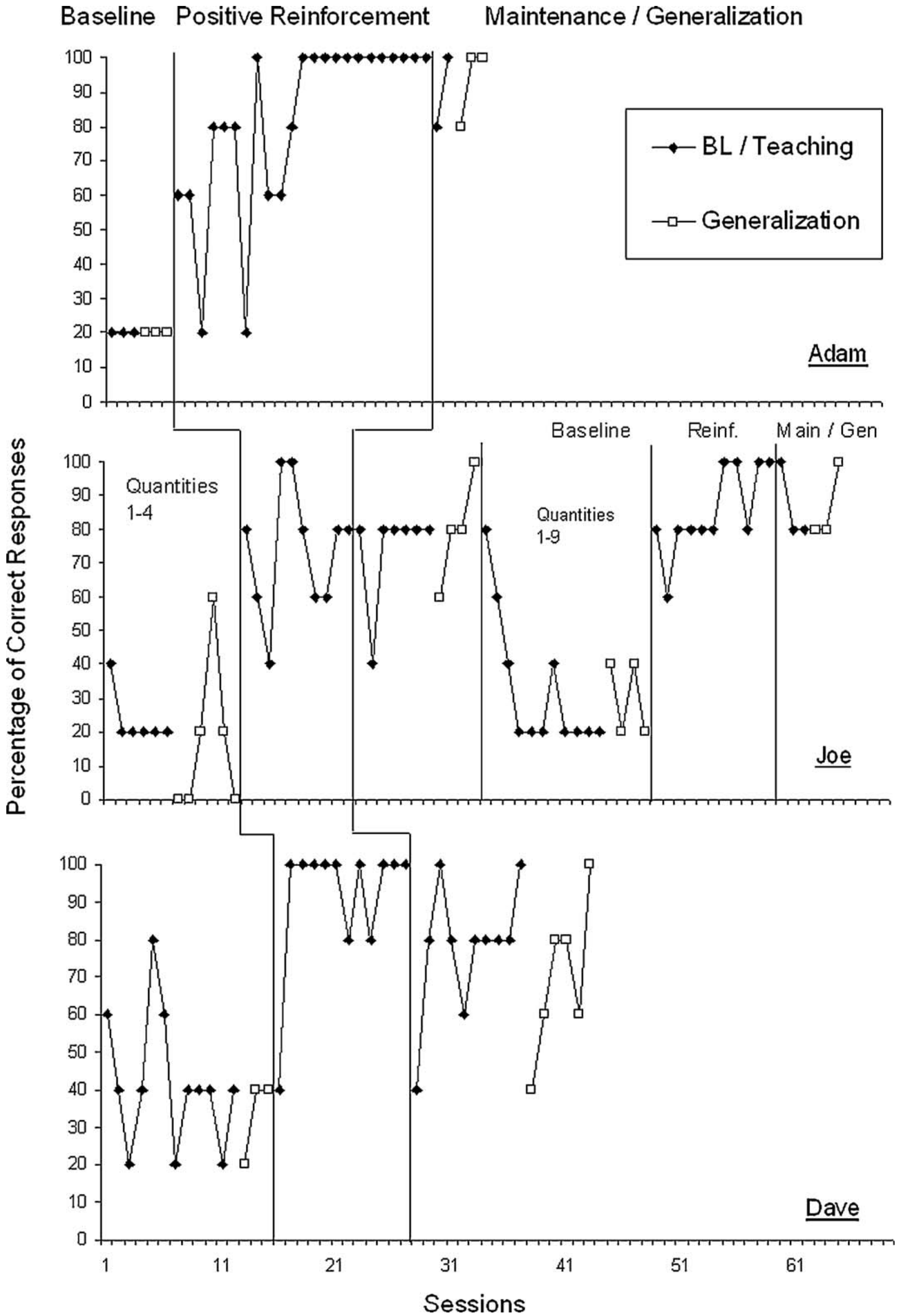

Fig. 2. Percentage of correct responding across *baseline*, *positive reinforcement*, and *maintenance* phases for all participants. Black diamonds depict performance with stimuli that were directly taught during the *positive reinforcement* phase. White squares depict performance with stimuli that were never included in the *positive reinforcement* phase (i.e., generalization stimuli).

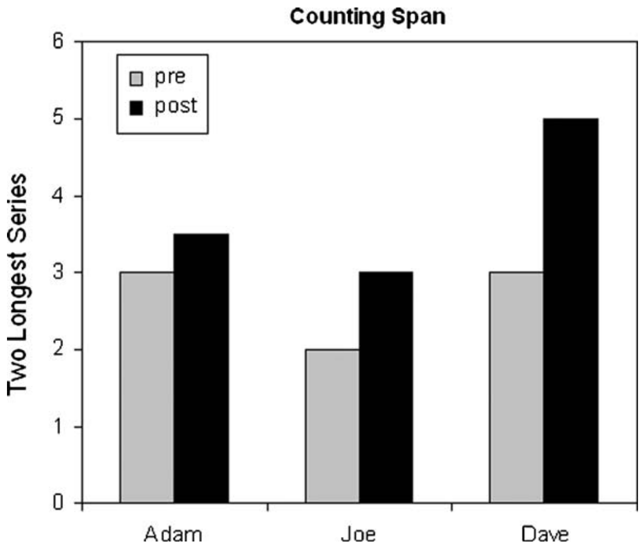

Fig. 3. Pre and post-test scores for all participants on the *Counting Span* subtest of the AGTB.

3. Discussion

The main focus of the present study was to answer two questions. First, whether performance on tasks that measure the central-executive of WM can be improved with the use of basic behavioral intervention procedures, and second, whether these effects generalize to novel stimuli which are not included in training. The results of the experiment demonstrate that positive reinforcement effectively improved central-executive WM abilities, at least as measured by CS tasks, in three children with autism. Further, these effects maintained after reinforcement was discontinued and were generalized to novel stimuli and behaviors (i.e., colors, shapes, and quantities counted). This is the first controlled study, of which the authors are aware, on the use of behavioral intervention procedures to improve EF or WM in children with ASDs.

The fact that WM performance did not decrease when positive reinforcement was discontinued is particularly encouraging. If only the *baseline* and *positive reinforcement* phases are examined, one potential interpretation of the findings is that performance was low in baseline simply because there was insufficient motivation to perform in baseline, due to a lack of feedback or reinforcement for responding. From this perspective, it would be possible that the *positive reinforcement* intervention did not actually improve participants' WM abilities, but rather simply gave them motivation to display the abilities they already possessed. However, when positive reinforcement was discontinued during the maintenance and generalization conditions (a return to baseline conditions), correct responding remained high, despite the fact that it no longer produced reinforcement. This pattern of results suggests that the increase in responding observed during *positive reinforcement* was not likely due to an increase in motivation alone, and may have resulted in a real and enduring improvement in the actual skill displayed during the WM task.

A secondary analysis of pre- and post-scores on the AGTB suggests that the positive reinforcement intervention may have resulted in increases on AGTB scores as well. However, since the AGTB scores were not evaluated in the context of an experimental design, future research utilizing a proper design will be needed before definitive conclusions can be made regarding the effects of the intervention on AGTB performance.

Some speculation regarding collateral behaviors which may have mediated WM performance warrants discussion. Experimenters repeatedly made the anecdotal observation that, during the *positive reinforcement* condition, participants engaged in higher rates of potential mediating behaviors, such as self-talk, counting on fingers, etc. No data were collected on these behaviors but it is an interesting finding because no such strategies or mediating behaviors were ever discussed with the participants or modeled to them in any way. It is possible that these behaviors were already in participants' repertoires, and that when participants engaged in them and then received reinforcement for correct responding, that these behaviors were also reinforced because they helped facilitate the correct response that immediately followed them. That is, an indirect reinforcement effect may have been observed on these responses, since these responses helped increase accurate responding to the CS task. However, such possibilities are pure speculation since no data were collected on the occurrence of these behaviors and no attempt was made to evaluate their influence, if any, on correct task performance. Future research on behavioral intervention strategies for EF and WM should include such measures.

One potential limitation of the current study is that generalization was not examined in a broader manner. Participants' WM was not evaluated in other settings nor did we evaluate whether improvements in WM were apparent in the course of participants' everyday activities. Moreover, no attempt was made to assess whether the intervention produced a clinically useful improvement. Although the lack of data on broader generalization is a limitation to the current study, the attainment of such generalization was not the purpose of the study. The main goal was to examine whether an effect could be attained on WM tasks and whether any generalization could be produced at all. That is, whether behavioral procedures can produce any improvement and, if so, to ensure that the demonstrated improvement was not simply rote memorization of particular behaviors. The results of the current study suggest a positive answer to both points.

Another potential limitation of the current study is the fact that the quantity 5 was included in both the training and generalization stimulus sets. Therefore, correct responding to the quantity 5 during generalization probes does not represent generalization across quantities. However, the quantity 5 was never included more than once per trial, thereby responding to this quantity never represented more than 25% (for Joe and Adam) or 20% (for Dave) of the correct responding observed. That is, it would not have been possible for participants to score above 25% correct on generalization probes if no generalization across quantities had actually occurred. Furthermore, generalization stimuli always included untrained colors and untrained shapes, thereby still demonstrating stimulus generalization, regardless of whether the quantity 5 was included.

The results of this study provide initial evidence which suggests that behavioral principles of learning and motivation may play an important role in EF and WM. It should be noted that adopting a behavioral perspective on WM does not deny the contribution of the physiology and neurology of the individual. Neurological research is critical because it identifies the neurological events which participate in the individual's interaction with their environment during episodes of working memory. But working memory also always involves an organism interacting with its environment, that is, there is always behavior occurring in relation to events in the environment. Therefore, the relations between behavior and environment must be investigated as well. Interdisciplinary collaborations between neuroscientists and behavior analysts, involving investigation of the relations between behavior, environment, and neuronal events, under conditions labeled as "executive functioning" and "working memory," would likely be fruitful and future investigators are encouraged to undertake such work.

This study is among the first to demonstrate a significant treatment effect on a measure of EF in general or WM in particular, in individuals with ASD. Much future research is still needed on a broader range of skills relevant to EF and WM and on the generalization of treatment gains across participants' everyday lives. Future research should also attempt to evaluate procedures which address a broader range of WM performances, such as abilities which are said to depend on the visuospatial sketchpad, the phonological loop, and others. Further research on behavioral intervention in these areas is likely to expand treatment options, as well as expanding the horizon of psychological phenomena addressed within the behavioral field.

References

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuo-spatial short-term and working memory in children: Are they separable? *Child Development, 77*, 1698–1716.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (DSM-IV)* (4th ed.). Washington, DC: APA.

Andersson, U. (2008). Working memory as a predictor of written arithmetical skills in children: The importance of central executive functions. *British Journal of Educational Psychology, 78*, 181–203.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science, 4*, 417–423.

Bennetto, L., Pennington, B. F., & Rogers, S. J. (1996). Intact and impaired memory functions in autism. *Child Development, 67*, 1816–1835.

Bull, R., Johnson, R. S., & Roy, J. A. (1999). Exploring the role of the visio-spatial sketchpad and central-executive in children's arithmetical skills: Views from cognition and developmental neuropsychology. *Developmental Neuropsychology, 15*, 421–442.

Catania, A. C. (1998). *Learning*. Upper Saddle River, NJ: Prentice Hall.

Cuvo, A. J. (1974). Incentive level influence on overt rehearsal and free recall as a function of age. *Journal of Experimental Child Psychology, 18*, 167–181.

Daffner, K. R., & Searl, M. M. (2008). The dysexecutive syndromes. In M. J. Aminoff, F. Boller, & D. F. Swaab (Eds.), *Handbook of clinical neurology neuropsychology and behavioral neurology* (pp. 249–). Amsterdam: Elsevier. [Chapter 12].

Dawson, P., & Guare, R. (2004). *Executive skills in children and adolescents: A practical guide to assessment and intervention*. New York, NY: The Guilford Press.

Farb, J., & Throne, J. (1978). Improving the generalized mnemonic performance of a Down syndrome child. *Journal of Applied Behavior Analysis, 11*, 413–419.

Fournet, N., Morand, O., Roulin, J. L., Naegele, B., & Pellat, J. (1996). Working Memory in mediated patients with Parkinson's disease: The central executive seems to work. *Journal of Neurological Neurosurgical Psychiatry, 60*, 313–317.

Geurts, H., Verte, S., Oosterlaan, J., Roeyers, H., & Sergeant, J. (2004). How specific are executive functioning deficits in attention deficit hyperactivity disorder and autism? *Journal of Child Psychology & Psychiatry, 45*, 836–854.

Granpeesheh, D., Tarbox, J., & Dixon, D. (2009). Applied behavior analytic interventions for children with autism: A description and review of treatment research. *Annals of Clinical Psychiatry, 21*, 162–173.

Hasselhorn, M., Schumann-Hengsteler, R., Gronauer, J., Grube, D., Mähler, C., Schmid, I., et al. (in press). *Arbeitsgedächtnistestbatterie für Kinder von 5 bis 12 Jahren (AGTB 5-12)*. Göttingen: Hogrefe-Verlag.

Hayes, S. C., Gifford, E. V., & Ruckstuhl, L. E. (1996). Relational frame theory and executive function: A behavioral approach. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive function*. Baltimore, MD: Brookes.

Hill, E. L. (2004). Evaluating the theory of executive dysfunction in autism. *Developmental Review, 24*, 189–233.

Hughes, C. (2001). Executive dysfunctions in autism: Its nature and implications for the everyday problems experienced by individuals with autism. In J. Burack & T. Charman (Eds.), *The development of autism: Perspectives from theory and research* (pp. 255–275). New York: Erlbaum.

Loftus, G. R. (1972). Eye fixation and recognition memory for pictures. *Cognitive Psychology, 3*, 525–551.

Ozonoff, S. (1995a). Reliability and validity of the Wisconsin Card Sorting Test in studies of autism. *Neuropsychology, 9*, 491–500.

Ozonoff, S. (1995b). EFs in autism. In E. Schopler & G. Mesibov (Eds.), *Learning and cognition in autism* (pp. 199–219). New York: Plenum.

Ozonoff, S. (1997). Components of executive function in autism and other disorders. In J. Russell (Ed.), *Autism as an executive disorder* (pp. 179–211). Oxford: Oxford University Press.

Ozonoff, S., & Jensen, J. (1999). Specific EF profiles in three neurodevelopmental disorders. *Journal of Autism and Developmental Disorders, 29*, 171–177.

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry, 32*, 1081–1105.

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry, 37*, 51–87.

Prior, M., & Hoffman, W. (1990). Neuropsychological testing of autistic children through exploration with frontal lobe tests. *Journal of Autism and Developmental Disorders, 20*, 581–590.

Ruble, L., & Scott, M. (2002). Executive functions and the natural habitat behaviors of children with autism. *Autism, 6*, 365–381.

Rumsey, J., & Hamburger, S. (1988). Neuropsychological findings in high functioning men with infantile autism, residual state. *Journal of Clinical and Experimental Neuropsychology, 10*, 201–221.

Skinner, B. F. (1974). *About behaviorism*. New York: Vintage Books.
Tamm, L., Hughes, C., Ames, L., Pickering, J., Silver, C. H., Stavinoha, P., et al. (2009). Attention training for school-aged children with ADHD: Results of an open trial. *Journal of Attention Disorders* [Epub ahead of print, PMID: 19805621].
Tervo, R. (2003). Identifying patterns of developmental delays can help diagnose neurodevelopmental disorders. *A Pediatric Perspective, 12*, 1–6.
Turley-Ames, K. J., & Whitfield, M. M. (2003). Strategy training and working memory task performance. *Journal of Memory and Language, 49*, 446–468.
Verté, S., Geurts, H. M., Roeyers, H., Oosterlaan, J., & Sergeant, J. A. (2006). The relationship of working memory, inhibition, and response variability in child psychopathology. *Journal of Neuroscience Methods, 151*, 5–14.