ELSEVIER

Volume 5, Issue 2, April–June 2011

ISSN 1750-9467

Research in Autism Spectrum Disorders

Research in Autism Spectrum Disorders 5 (2011) 855–863

Contents lists available at ScienceDirect

Research in Autism Spectrum Disorders

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

Further analysis of the effects of positive reinforcement on working memory in children with autism

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

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

ARTICLE INFO

Article history:
Received 23 September 2010
Accepted 27 September 2010

Keywords:
Autism
Executive functions
Working memory
Positive reinforcement
Central executive

ABSTRACT

Individuals with autism spectrum disorders (ASD) often exhibit impaired executive function (EF) performance, including difficulty with working memory (WM), in particular. While research has documented the existence of these deficits, surprisingly little research exists that evaluates potential treatment strategies for improving EF or WM. One exception is a study that used positive reinforcement to improve performance on a classical WM task, the counting span, resulting in both maintenance and generalization (Baltruschat et al., 2011). The current study is the second in a programmatic line of research on behavioral intervention for improving WM in children with autism. This study extended the use of the same procedure (positive reinforcement) to another task which is said to measure WM, a Complex Span, and included three additional children with autism. Results demonstrated significant improvements in performance for each participant, including maintenance and generalization to untrained stimuli and untrained responses. These results provide further evidence that behavioral intervention procedures may be useful for improving skills labeled as EF or WM in children with ASD.

Individuals with autism spectrum disorders (ASD) suffer from impairment in communication and social interaction as well as demonstrate restricted, repetitive and stereotypical patterns of behavior and interests (American Psychiatric Association; APA, 1994). Delays in areas such as cognition, play, daily living skills, and motor development are also common (Tervo, 2003). Impairment in executive function is an additional area of concern with many children with ASD (Hughes, 2001; Ozonoff, 1995; Ozonoff & Jensen, 1999). The term executive function (EF) generally refers to a collection of presumably cognitive processes involved in such things as planning, goal persistency, cognitive flexibility, abstract thinking, rule acquisition, initiating appropriate actions, inhibiting inappropriate actions, and selecting relevant sensory information (Hill, 2004). Given the broad and complex abilities to which the term EF refers, it is not surprising that it is somewhat ill-defined, with no broad-based consensus regarding a definition or model of EF between as well as within various scientific disciplines (Ruble & Scott, 2002).

* 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 © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.rasd.2010.09.015

Working memory (WM) is a commonly researched component of EF and is said to refer to one's ability to "keep information online" while simultaneously processing it. Most agree that the WM construct consists of several different components, however, there is little agreement on the exact nature and composition of the components (Alloway, Gathercole, & Pickering, 2006). A detailed model of WM describing separate memory systems responsible for the temporary storage of visuospatial and phonological representations was postulated in 1974 by Baddeley and Hitch (Baddeley, 2000). The visuospatial sketchpad, the phonological loop, and the central executive are described as the three main WM components of their system. According to Baddeley, visual information is processed in the visuospatial sketchpad, the phonological loop is responsible for auditory information, and the central executive represents a supervisory system, controlling the flow of information from the other two systems. In 2000, a fourth system, the "episodic buffer," was added, which is said to link visual and auditory information within the correct chronological sequence.

Research on WM in children with ASD has documented deficits across a range of ages and functioning levels (Geurts, Verté, Oosterlaan, Roeyers, & Sergeant, 2004; Ozonoff, 1997; Verté, Geurts, Roeyers, Oosterlaan, & Sergeant, 2006; see Hill, 2004 for a recent review). The central executive component of WM in particular has been shown to be impaired in children with ASD (Ozonoff, Pennington & Rogers, 1991; Prior & Hoffman, 1990; Rumsey & Hamburger, 1988). Shifting between tasks, retrieving new strategies, inhibiting inappropriate reactions, and strengthening selective attention are all included as tasks of the central executive.

Like most cognitive constructs, the presence of WM (and of the central executive in particular) is inferred by measuring performance on tasks that involve overtly observable behavior. One such set of tasks is referred to as "complex span tasks" (Andersson, 2008). Complex span tasks are based on a "dual-task paradigm," combining a memory span measure with a concurrent processing task. They are said to measure a participant's ability to simultaneously store and process information and were developed to prevent the participant from rehearsing the stimuli in order to memorize it (Bull, Johnson, & Roy, 1999; Fournet, Moreand, Roulin, Naegele, & Pellat, 1996). An early complex span task is the so called "reading span," first invented by Daneman and Carpenter in 1980. In reading span tasks, a number of sentences are presented to the participant while the participant is required to memorize the last word of the sentence. The words have to be recalled in the correct order as soon as the last sentence was presented. Each time a sentence is presented, the participant must simultaneously remember the last word of the previous sentence and identify and remember the last word of the current sentence, therefore uninterrupted rehearsal is difficult or impossible (Swanson & Ashbaker, 2000). Another variation of the complex span task, including classification responses as the distracter task, was first developed and experimentally tested by Zoelch, Seitz, and Schumann-Hengsteler (2005). This task involves the presentation of a sequence of visual stimuli and for each stimulus, the participant is asked to emit a classification response according to the function of the object (e.g., "Can you eat it?"). At the end of the sequence of the stimuli, the participant is asked to state the names of the stimuli in the order they had been presented.

While a significant amount of research has documented the presence of EF and WM impairment in individuals with ASD, surprisingly little research has been done on methods for improving it. The small amount of treatment research that exists generally focuses on children with attention deficit hyperactivity disorder (ADHD), fetal alcohol spectrum disorder, or Down syndrome. An early study from 1978 showed that a rehearsal training program effectively improved the mnemonic performance of a child with Down syndrome (Farb & Throne, 1978). Another utilizing rehearsal increased WM span scores in typically developing children (Turley-Ames & Whitfield, 2003). A recent study showed that the performance on neuropsychological tests in EF areas, including WM, increased in children with ADHD by training their selective, alternating, and divided attention (Tamm et al., 2010).

Cognitive constructs in general, and EF and WM in particular, tend to receive little attention in behavior analytic research. Basic researchers conducted a limited number of studies demonstrating that positive reinforcement can affect performance on short-term memory tests with typically developing adults in the 1970's (e.g., Cuvo, 1974; Loftus, 1972) but little work has been published since. Further, little behavioral research has addressed WM, and virtually no research is being done on the application of behavioral principles to WM in clinical populations. It seems likely that this dearth of behavioral research on cognition is partially responsible for the common but false belief that cognitive events cannot (or should not) be studied by behavior analysts. Although this belief remains a common one, it is misguided. Starting in 1945, Skinner described the basic assumptions which forged the philosophical foundation for contemporary behavior analytic psychology, that is, Radical Behaviorism (Skinner, 1945). The core of Radical Behaviorism is the assumption that anything anyone does in interacting with their environment is to be considered behavior, be it overt or covert. Furthermore, all behavior is assumed to be subject to the same basic principles of learning and motivation (i.e., reinforcement, extinction, stimulus control, punishment), regardless of whether it is covert or overt. Thus, mental or cognitive events, insofar as they refer to something a person is actually doing, are not mental or cognitive at all, but are rather "private" behaviors. The only defining difference between public and private events is that public events are amenable to observation by more than one person simultaneously, whereas private events are not (Skinner, 1974).

Approaching WM from a radical behavioral perspective is inherently practical because decades of research on the principles of behavior are already available, all of which is assumed to apply equally to private behavior. This literature should be a rich source of ideas for how research and practical work in the area of WM may proceed. 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, it might be pointed out that individuals responding to tests of WM are almost always doing something in addition to merely providing the correct or incorrect answer. That is, there are other behaviors occurring which may affect the probability of the correct response occurring. If such "mediating" behaviors, such as rehearsal, play a critical role in working memory, then the behavioral literature should have something to say about how to teach and maintain those behaviors. For example, the basic behavioral intervention procedures of prompting and reinforcement may be useful for establishing these mediating behaviors in people who do not have them.

Despite the apparent utility of a behavioral approach to improving WM, very little previous research has been done on the topic. One exception is a programmatic line of research on behavioral intervention for WM that our group has been conducting. In our first study on the topic, we demonstrated that the use of positive reinforcement produced a significant improvement in the WM performance of three children with autism (Baltruschat et al., 2011). A counting span task, a common procedure for investigating the performance of the central executive component of working memory, was selected. In this task, participants were presented with a series of visual stimuli, consisting of quantities of shapes on flashcards, and were required to count the quantities on each card and then later recall the quantities in the correct order in which the flashcards were presented. Results demonstrated clear and significant increases in accuracy among all three participants. Furthermore, results maintained after reinforcement was discontinued and generalization to a task involving untrained quantities and untrained shapes was obtained.

The current study extended the use of positive reinforcement to another WM task, a complex span task, and included three additional children with autism. In contrast to the counting span task, the complex span task is said to measure central executive abilities which are supported by the phonological subsystem of working memory, whereas the counting span task is said to assess central executive skills supported by the visuospatial subsystem. From a radical behavioral perspective, WM tasks are not assumed to measure the performance of anything else – they are assumed to be legitimate performances in their own right. However, if a behavioral approach to improving WM is to be meaningful, it must be demonstrated to be effective across a broad range of WM tasks, not merely one or two. Hence the need for replication across additional behaviors which are said to be tests of working memory. Furthermore, it must be demonstrated to produce stimulus and response generalization, not merely proficiency with tasks that were directly trained.

The purpose of this study was to investigate the effects of positive reinforcement on an entirely different task (i.e., complex span), which still addresses the area of functioning referred to as EF and WM by the general psychology community. Furthermore, a more complex task was included in the current study. Specifically, the distracter task in the counting span involves counting quantities of visual stimuli, a relatively straightforward performance. Particularly for someone who has a significant reinforcement history for counting, thereby establishing a strong intraverbal linkage between saying numbers in chronological order (Skinner, 1957), the performance may not be excessively demanding. However, the distracter task included in the complex span task in the current study was more complex (i.e., classification). A full conceptual treatment of classification behavior is beyond the scope of this paper, but classification likely involves responding to the relation between stimuli – referred to as "hierarchical relational responding" (Hayes, Barnes-Holmes, & Roche, 2001).

In summary, the current study further investigated the effectiveness of positive reinforcement for improving performance on tasks that are reported to measure WM, and attempted to extend previous findings by including a different (and more complex) task, and by including three other children with ASD. The goal of the study was to demonstrate a generalized improvement in the task, not merely rote memorization of particular stimuli, as well as maintenance after treatment was withdrawn.

1. Method

1.1. Participants and setting

Three children were recruited for this study from the clients of a community-based provider of home-based behavioral intervention services. The criteria for participation in the study were as follows: (1) age of six years or greater, (2) an ASD diagnosis, (3) relatively well-developed language repertoire (ability to follow novel complex instructions, etc.), and (4) the child's family and therapy team had to express that improvement in the general area of executive function – and working memory in particular – was a clinical priority. Before the study began, the procedures received approval by an institutional review board and all participants' parents gave written informed consent.

Ken was a 6-year-old boy with autism. Prior to Ken's participation in this study, he had received approximately five years of home-based behavioral intervention. At the time of the study, Ken was receiving 18 h of behavioral intervention per week. Ken was in good physical health, with no history of seizures, accidents, or hospitalizations. Throughout the duration of the study, Ken was on a gluten and casein-free diet.

Larry was an 8-year-old boy with autism and Attention- Deficit/Hyperactivity Disorder. Prior to Larry's participation in the study, he had received seven years of home-based behavioral intervention. At the time of the study, he was receiving 13 h per week of behavioral intervention. Larry was in good physical health, with no history of seizures, accidents, or hospitalizations.

Alex was an 8-year-old boy with autism. Prior to Alex's participation in this study, he had received approximately four years of behavioral intervention. At the time of the study, Alex was receiving 10 h per week of behavioral intervention. Alex displayed good physical health, with no history of seizures, accidents, or hospitalizations.

Sessions during all phases other than the AGTB were conducted in the participants' home environment, where their behavioral intervention sessions normally took place (usually their bedroom or a 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. Response measurement and interobserver agreement

Data were collected on the accuracy of responding for each trial, during sessions of all phases other than the AGTB. A correct response was defined as the participant accurately stating the names of the stimuli 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 74%, 62%, and 100% of sessions for Ken, Larry, and Alex, 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 87% (range = 60–100%), 86% (range = 60–100%), and 98% (range = 80–100%) for Ken, Larry, and Alex, respectively.

1.3. Pre- and post-training WM assessment: Arbeitsgedaechtnis Testbatterie

At the outset and at the very end of the study, all participants completed the Arbeitsgedaechtnis Testbatterie (AGTB), translated as "Working Memory Test Battery" (Hasselhorn et al., in press). The AGTB is a computerized testing protocol that uses a touch-sensitive monitor and has been translated into English for use with English-speaking participants. It was administered in order to collect detailed performance information regarding each child's central executive WM abilities, and, more specifically, to identify if children displayed difficulty in the complex span subtest. The central executive AGTB scale comprises six subtests, including: (1) *Complex Span*, (2) *Color Span*, (3) *Digit Span Backwards*, (4) *Stroop-Like*, (5) *Go/No Go*, and (6) *Counting Span*. These subscales have been described elsewhere (Baltruschat et al., 2011). The *Complex Span* task is the most relevant to the current study. In this task, a sequence of visual stimuli are presented to the participant. In response to each stimulus, participants were asked to provide a classification response that identified a function of the object (e.g., Can you eat it?). At the conclusion of the sequence, the participants were then asked to list the stimuli in the order in which they were initially presented. All AGTB sessions were conducted in a treatment room within 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.

1.4. Materials

Fig. 1 displays the material used during all phases except for the AGTB. The materials were developed in order to measure the same skill/ability as its equivalent subtest (*Complex Span*) of the AGTB. However, the material differed from its original AGTB task, in that they included novel pictures. The materials consisted of 8 flashcards showing pictures of two-syllable words, half of them showing edible items and half of them non-edible ones. Flashcards were presented in random order. After each flashcard, a distracter question requiring a classification response (i.e., "can you eat it/can you wear it") was provided by the investigator and the participant was required to answer with "yes" or "no" before the next flashcard was presented to the participants. After the entire sequence of cards had been presented, the experimenter asked the participant what he remembered in the correct order. The 8 flashcards used during the generalization phase were evenly split in pictures showing items that you can wear and items that you cannot.

1.5. Procedures

1.5.1. Pre-baseline evaluation

Before collecting baseline data, a brief pre-baseline evaluation was conducted in order to identify the level of difficulty set for baseline. To determine baseline difficulty, the number of items a child could successfully memorize (without error), for two consecutive trials, was probed. Based on the results, each participant's training level, which would be used during the actual intervention, was determined and evaluated. This process began with presenting a sequence of two flashcards. If the participant responded correctly for two consecutive trials, the number of flashcards was increased by one on the next trial. If the participant again responded correctly for two consecutive trials, another card was added during the next trial. This procedure was continued until an error was made on two consecutive trials. This number of flashcards was then selected for inclusion in baseline and training (initially 2 for Ken, 5 for Larry, and 5 for Alex).

1.5.2. Baseline

Prior to the first flashcard's presentation in baseline, the investigator provided instructions to the participant. Also, the investigator twice rehearsed the task with the child before beginning the session. Trials were then presented in the same manner as in the *pre-baseline evaluation*, except that the target number of stimuli were always presented and sessions consisted of five trials. 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.

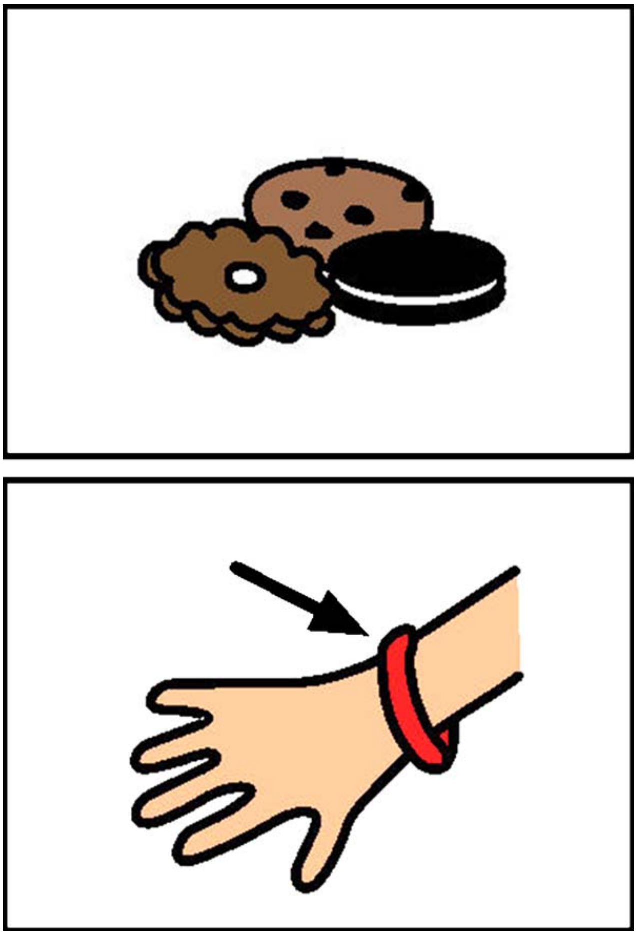

Fig. 1. Sample flashcards used during *baseline* (cookies) and *generalization* (bracelet) phases.

1.5.3. Positive reinforcement

Sessions were identical to *baseline,* with the following exceptions. Prior to the beginning of each session in this condition, the participant selected an item from an array that they would earn during that session. All items included in the array were ones which participants did not have access to outside of experimental sessions. Contingent on each correct response to a trial, the participant received access to the item for 1 min or, in the case of edibles, were allowed to consume a small amount of the item. If a participant responded incorrectly to a trial, corrective feedback was given (e.g., a neutral "no, that's not quite right") and the next trial was conducted. At the outset of each session, the reinforcement contingencies were explained to the participant. Participants were given a five min break between sessions.

1.5.4. Reinforcement and verbal prompting

If the *positive reinforcement* condition did not result in a large and stable increase in accuracy, a verbal prompt procedure was added. In this phase, the child was prompted to vocally rehearse the stimuli presented cumulatively and overtly in unison with the experimenter. That is, each stimulus presented to the child had to be loudly repeated four times while the stimulus was visually present and then one more time during the inter-stimulus interval. After a large and stable increase in accuracy was produced, the experimenter's voice was gradually faded out, then the experimenter instructed the participant to decrease their own volume, gradually from talking aloud, to whispering, to only moving lips, and finally to covert rehearsal.

For one participant (Larry) it was hypothesized that the rehearsal strategy could be more easily learned if the task was made easier. Consequently, a difficulty fading procedure for the amount of stimuli presented was also included. The number of flashcards presented during a trial was decreased from 5 to 2. The number of flashcards presented per trial was gradually increased to 5, contingent on continued correct responding.

1.5.5. Reinforcement and visual prompting

If the *reinforcement and verbal prompting* phase did not produce a large and stable increase in accuracy, a visual prompting procedure was added. During this treatment phase, the picture stimuli remained visible on the table throughout an entire trial while the child engaged in the overt cumulative rehearsal behavior. After a large and stable increase in accuracy was produced, the number of flashcards remaining visible on the table were gradually faded out from all five cards, to 4, then 3, and ultimately to no cards on the table. When one card was removed, a white piece of paper with a black outline was put in its place. The outline stimuli were faded one step behind the presence of the picture cards. That is, when the second flashcard was removed from the table, the outline of the first flashcard was removed. Then, when the third flashcard was removed from the table, the outline of the second card was removed as well, and so on. Finally, when all five flashcards were removed from the table the outlines of the last two flashcards were still visible. The experimenter then proceeded to remove the outlines completely.

1.5.6. Maintenance

When participants showed large and stable increases in accuracy, relative to baseline, the effective treatment was removed and the *maintenance* phase was initiated. This phase was identical to *baseline.*

1.5.7. Generalization

When participants showed consistent and stable performance in the maintenance phase, a *generalization* phase was introduced. The *generalization* phase consisted of introducing trials with stimuli which had never been present during training and were asked distracter questions which required untrained categorization responses ("Can you wear it" during generalization, versus "Can you eat it?" during the other phases).

2. Results

Fig. 2 depicts the percentage of correct responding for trained and generalization tasks across all phases, for all participants. The top panel shows Ken's data. During the *baseline* phase for Ken, correct responding was consistently low (mean = 20%) and remained low in response to generalization stimuli (mean = 14%). When the *positive reinforcement* phase

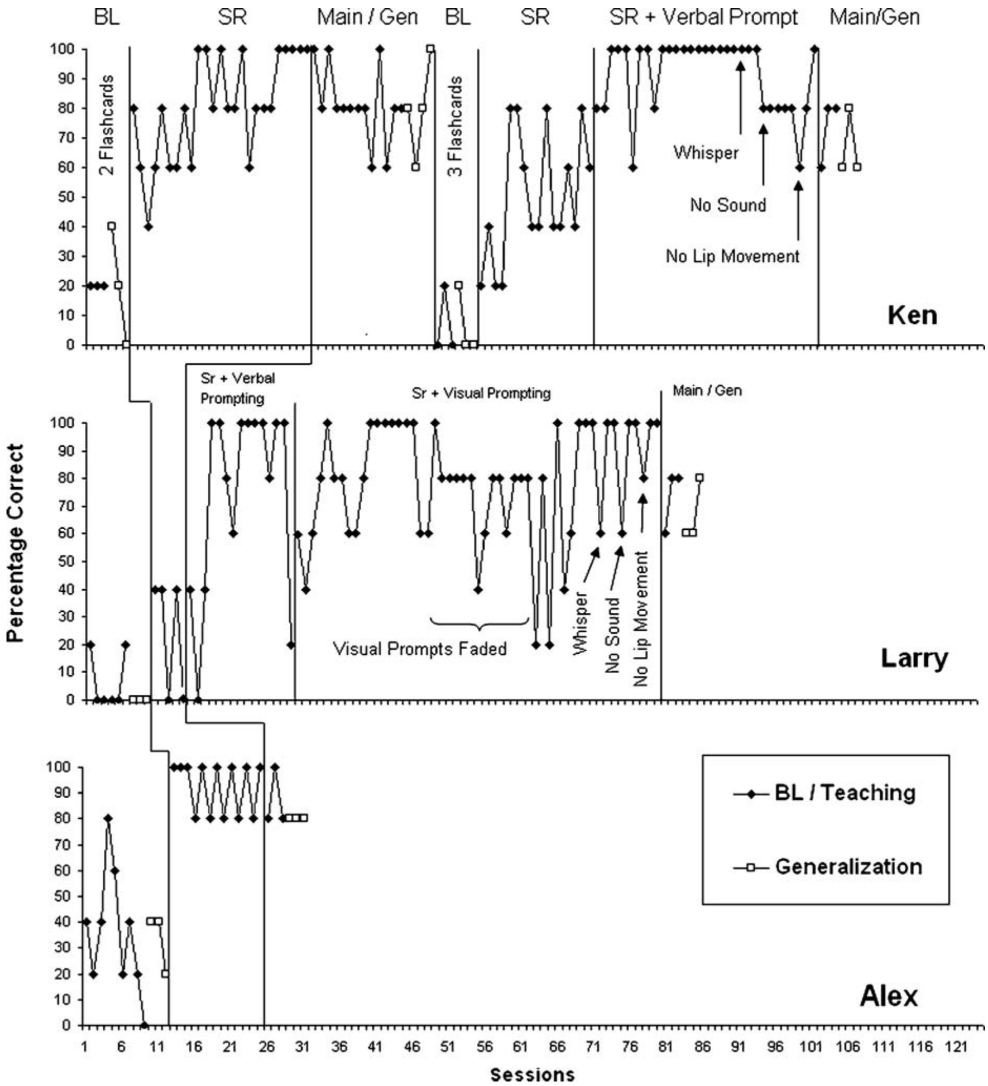

Fig. 2. The percentage of correct responding across phases for all participants. Black diamonds depict responding to stimuli that were directly taught. White squares represent accuracy during generalization probes with untrained stimuli and responses.

was initiated, Ken's correct responding increased immediately and continued to demonstrate a gradual increase until it remained stable at 100% for 5 consecutive sessions (mean = 81%). The maintenance phase was then initiated, wherein reinforcement was discontinued. Ken continued to respond correctly with high accuracy (mean = 82%). He also demonstrated accurate responding to generalization stimuli, which had never been directly taught or reinforced (mean = 80%).

Due to Ken's rapid response to the intervention, it was decided to increase the level of difficulty of the task by adding a flashcard to the sequence, resulting in sequences of 3 flashcards. During the *baseline* phase with 3 flashcards, correct responding was consistently low (mean = 7%) and continued to be so in response to the generalization stimuli (mean = 7%). When the positive reinforcement phase was initiated, Ken's correct responding increased significantly, relative to baseline, however it did not consistently meet or exceed 80% (mean = 55%). Therefore the verbal prompting and reinforcement phase was introduced. Ken's responding immediately increased, until he was consistently responding 100% correct (mean = 92%). After 12 consecutive sessions at 100% correct, a prompt fading procedure was introduced in which the experimenter's voice was faded out and the volume of Ken's voice was decreased. Throughout the fading procedures Ken responded correctly between 60% and 100% of the time. He continued to respond correctly during maintenance (mean = 73%) and demonstrated significantly better accuracy during generalization (mean = 67%), relative to baseline.

The middle panel depicts Larry's data. During the initial *baseline* phase, Larry demonstrated consistently low levels of correct responding (mean = 7%). His responses remained low with the generalization material as well (mean = 0%). When *positive reinforcement* was introduced, Larry's accuracy increased slightly but the improvement was not sufficient (mean = 24%). Consequently, the verbal prompting and reinforcement procedure was introduced, while simultaneously reducing the number of flashcards in the sequence from 5 to 2. Larry's accuracy increased significantly and remained high while the number of flashcards was gradually faded back up to 4 (indicated by arrows on the graph). However, when the number of cards was increased to 5, his performance decreased to 20%. At this time, a third phase was introduced, involving a visual prompting procedure. This phase resulted in a gradual increase in accuracy to 100%. After 7 consecutive sessions at 100% correct, a prompt fading procedure, for both the verbal and the visual prompt, was introduced. After all prompts were successfully removed, the maintenance phase was initiated. Larry continued to respond correctly between 60% and 80% of the time during maintenance (mean = 73%). He also demonstrated higher accuracy with the generalization tasks (mean = 67%) than before intervention (mean = 0%).

The bottom panel depicts Alex's data. During baseline, Alex showed high variability in responding across training and generalization stimuli (mean = 36% and 33%, respectively). He then demonstrated an immediate increase during the positive reinforcement phase (mean = 92%), and maintained correct responding (mean = 87%), as well as demonstrating generalization (mean = 80%), during the maintenance phase.

2.1. Secondary analysis: AGTB results

Fig. 3 shows the pre- and post-test scores of all three participants in the Complex Span subtest of the AGTB. The *y*-axis shows the mean number of stimuli in the two longest sequences of stimuli that were answered correctly for each participant. Ken's performance improved from 1.5 cards (before treatment) to 3 cards (after treatment). Alex's performance improved from 2.5 prior to treatment to 4.5 after treatment. Larry's accuracy changed from 3 before treatment to 3.5 after treatment.

3. Discussion

The complex span performance of all participants significantly improved as a result of intervention. These data support the possibility that skills labeled as EF and WM may be amenable to improvement via basic behavioral procedures, such as prompting and reinforcement. This study extends previous research (Baltruschat et al., 2011) by employing a different and more complex task and by replicating across three additional children with autism.

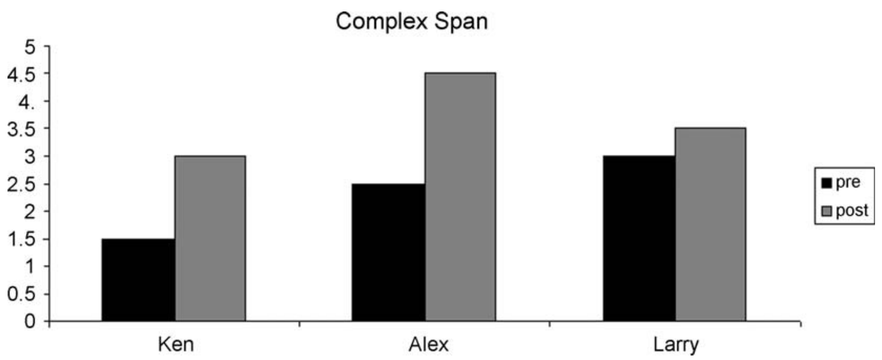

Fig. 3. Pre- and post-test scores for all participants on the Complex Span subtest of the AGTB. The *y*-axis depicts the mean of the two longest series that were correctly recalled for each participant.

Positive reinforcement was effective when implemented alone, for two of three participants (Alex and Ken). These results are encouraging because positive reinforcement is the most fundamental of behavioral procedures, can be simple to implement, and is non-intrusive. However, positive reinforcement alone did not produce a significant increase in accuracy for Larry, nor was it effective when the difficulty of the task was increased from a series of two stimuli to three stimuli for Ken. In both cases, explicit prompting and reinforcement of rehearsal was required to bring accuracy up to consistently high levels. These results suggest that mediating behaviors of some kind are likely necessary when WM tasks reach sufficient levels of complexity, and further, that these behaviors appear to be amenable to direct teaching. The threshold of difficulty of a given task that might require explicit teaching of mediating behaviors for a given participant is likely influenced by the degree to which the mediating behavior is already in the participant's repertoire, as well as the age and functioning level of the child. Future research should attempt to directly investigate these variables.

The generalization of accurate responding to an untrained categorization response (i.e., "Can you eat it?" as opposed to "Can you wear it?") with untrained stimuli is encouraging. Indeed, without such a finding, one might legitimately say that participants were merely trained to rotely memorize particular sequences of stimuli – an outcome of questionable significance. But the generalization demonstrated in this study shows that participants' overall ability to perform on complex span tasks was improved, a finding that further suggests that WM is susceptible to improvement via behavioral learning principles.

The finding that accurate responding maintained after positive reinforcement was discontinued is also positive. It would not be unreasonable to expect some degree of motivation to be necessary for participants to perform accurately on WM tasks but the continued need for contrived reinforcement is clearly less desirable in terms of clinical utility and everyday efficiency. The demonstration of maintenance also further suggests that the intervention procedures may have affected participants' overall ability to perform on complex span tasks, not merely gave them motivation to do so during the intervention phase. Although maintenance (as evidenced by increased accuracy relative to baseline) was observed for all participants, it is interesting to note that accuracy during the maintenance phase was slightly lower in every case. It is possible that this small difference reflects the portion of accurate responding that was due to the motivation provided by positive reinforcement at the moment that the sessions were conducted, as opposed to the maintained effect of the strengthening of the skill overall. This possibility remains purely speculative but it would be an interesting detail to investigate in future research.

As in our previous study on improving WM performance in children with ASD (Baltruschat et al., 2011), experimenters in the current study anecdotally reported that, even when a child was not prompted to engage in any type of rehearsal behavior, participants frequently appeared to be improvising their own mnemonic strategies, such as using fingers and self-talk. These mediating behaviors appeared to increase when positive reinforcement was implemented for correct responding to the tasks, even when no prompting of mediating behaviors was conducted. Unfortunately, no data were collected on these behaviors, nor did we evaluate whether such strategies were actually helpful, but it is an interesting finding because no strategies were discussed with the participants nor modeled to them in any way, in the *positive reinforcement* phase. Future research on behavioral intervention strategies for EF and WM should attempt to measure any mediating behaviors which may already be present in participant repertoires at the outset of the study.

One limitation of the current study is that generalization to participants' everyday lives was not evaluated. However, this study is one in a line of research that would be best considered bridge research, in that the demonstration of clinical utility is not the goal or purpose of the research. Rather, this line of research seeks to answer the question of whether, under controlled conditions, positive reinforcement and prompting can improve performance on WM tasks in individuals with ASD, and whether such improvement maintains and is not merely rote memorization. The results of the current study suggest a positive answer to these questions. Much more research will be needed which evaluates how to translate such findings into procedures which can be put into practice in clinical intervention settings.

The interplay of behavioral and neurological perspectives on WM may be worthy of discussion. Contrary to what many may believe, the adoption of a behavioral perspective on WM in no way denies the contribution of the organism's physiological and/or neurological constitution. Indeed, the importance of neurological research should not be underestimated, as it identifies the neurological events that contribute to the individual's interaction with their environment during WM episodes. But, WM involves more than just neurological events. WM, like every other psychological event, involves not only an organism equipped with specific neurological structures, but, equally as important, an organism involved in a reciprocal interaction with its environment. That is, there is always behavior occurring in relation to events in the environment. Therefore, the relationship between behavior and environment must be explored in all its facets and it is at this level that behavioral practitioners will eventually be able to provide effective treatments for WM deficits in individuals with ASD. It is hoped that future research will involve explicit collaborations between behavior analysts and researchers in neurology and cognitive neuroscience, as it is likely that the data with which these differing disciplines work would complement one another intellectually and practically.

This study is one of the first to demonstrate persuasively that, with regard to individuals with ASD, behavioral procedures can produce robust changes in tasks which are said to measure EF in general, and WM in particular. Furthermore, this study is among the first to provide the groundwork for what may eventually be demonstrated as an effective treatment for WM deficits in individuals with ASD. It is important that future research investigate the broader range of skills comprising EF and WM as well as the generalization of treatment gains across participant's every day life. Further, different measures of WM (e.g., stroop tasks, digit span backwards tasks, etc.) should be utilized, in order to broaden the scope of memory tasks to which behavioral strategies have been applied.

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.

Baltruschat, L., Hasselhorn, M., Tarbox, J., Dixon, D. R., Najdowski, A. C., Mullins, R. D., et al. (2011). Addressing working memory in children with autism through behavioral intervention. *Research in Autism Spectrum Disorders, 5*, 267–276.

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.

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.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior, 19*, 450–466.

Farb, J., & Throne, J. (1978). Improving the generalized mnemonic performance of a Down syndrome child. *Journal of Applied Behavioral 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.

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., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of language and cognition*. New York, NY: Kluwer Academic/ Plenum Publisher.

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. (1995). 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.

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. (1945). The operational analysis of psychological terms. *Psychological Review, 52*, 270–277.

Skinner, B. F. (1957). *Verbal behavior*. New York, NY: Prentice Hall.

Skinner, B. F. (1974). *About behaviorism*. New York, NY: Vintage Books.

Swanson, H. L., & Ashbaker, M. H. (2000). Working Memory and reading comprehension in learning disabled readers: Does the executive system have a role? *Intelligence, 28*, 1–30.

Tamm, L., Hughes, C., Ames, L., Pickering, J., Silver, C. H., Stavinoha, P., et al. (2010). Attention training for school-aged children with ADHD: Results of an open trial. *Journal of Attention Disorders, 14*, 86–94.

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.

Zoelch, C, Seitz, K., & Schumann-Hengsteler, R. (2005). From rag (bag)s to riches: Measuring the developing central executive. In W. Schneider, R. Schumann-Hengsteler, & B. Sodian (Eds.), *Young children's cognitive development: Interrelationships among executive functioning, working memory, verbal ability, and theory of mind*. Mahwah, NJ: Erlbaum.