Title: Inner speech is used to mediate short-term memory, but not planning, among intellectually high-functioning adults with autism spectrum disorder

Abstract

Evidence regarding the use of inner speech by individuals with autism spectrum disorder (ASD) is equivocal. To clarify this issue, the current study employed multiple techniques and tasks used across several previous studies. In Experiment 1, participants with and without ASD showed highly similar patterns and levels of serial recall for visually-presented stimuli. Both groups were significantly affected by the phonological similarity of items to be recalled, indicating that visual material was spontaneously recoded into a verbal form. Confirming that short-term memory is typically verbally mediated among the majority of people with ASD, recall performance among both groups declined substantially when inner speech use was prevented by the imposition of articulatory suppression during the presentation of stimuli. In Experiment 2, planning performance on a Tower of London task was substantially detrimentally affected by articulatory suppression among comparison participants, but not among participants with ASD. This suggests that planning is not verbally mediated in ASD. Importantly, the extent to which articulatory suppression affected planning among participants with ASD was uniquely associated with the degree of their observed and self-reported communication impairments. This confirms a link between interpersonal communication with others and intrapersonal communication with self as a means of higher-order problem-solving.

Keywords: Autism spectrum disorder, inner speech, verbal mediation, short-term memory, executive functioning

There is a long-standing debate in psychology and philosophy about the relation between language and thought (e.g., Carruthers, 2002; Sokolov, 1968/1972). Recent behavioural and neuropsychological studies have provided convincing evidence that several aspects of executive control depend to some extent on linguistic thinking (e.g., Baldo et al., 2005; Dunbar & Sussman, 1995; Gruber & Goschke, 2004). According to Vygotsky's influential (1934/1987) theory, the ability to "think in speech" (as opposed to visual imagery) is critical for flexible behaviour and cognition, and is the foundation for effective self-regulation. Crucially, Vygotsky argued that verbal thinking has its roots in linguistically-mediated exchanges with others (such as caregivers) early in life. These interpersonal dialogues, which serve as an external means of regulating the child's behaviour early in life, gradually become intrapersonal over time, such that the child is able to regulate their own behaviour by engaging in dialogue with self, in the absence of others. Initially, this self-talk is overt in the form of "private speech" (previously known as egocentric speech), which occurs almost universally among typically developing children (Winsler et al., 2003). Then, during middle childhood, self-talk becomes internalised to form "inner speech". Vygotsky viewed the conversion of private speech into inner speech as heralding the final shift from preverbal thought to fully intrapersonal verbal thinking.

In the current study, we explored the verbal mediation of different domains of cognition in autism spectrum disorder (ASD), a disorder of social-communication that, if Vygotskian theory is correct, should involve a significant diminution of inner speech use (see Fernyhough, 1996). Below, we outline evidence regarding the typical development of verbal mediation, before discussing existing evidence regarding verbal mediation among individuals with ASD.

The typical development of verbal mediation

Results from research involving typically developing children have been largely supportive of Vygotsky's (1934/1987) theory about the developmental course and functional significance of verbal thinking (e.g., Fernyhough & Fradley, 2005; Al-Namlah, Fernyhough, & Meins, 2006; see Winsler, Fernyhough, & Montero, 2009). In particular, research has focused on the development of inner speech use for the purpose of mediating short-term/working memory and executive functions.

In line with Vygotsky's (1934/1987) view that inner speech is not fully functional until middle childhood, several lines of evidence suggest that short-term memory is not fully verbally mediated until around six or seven years of age among typically developing children. In order to establish whether short-term memory for visually-presented information is verbally or visually mediated, studies have assessed the effect on serial recall of manipulations to the phonological (and visual) properties of the items to be recalled. Among typically developed adults, pictorial items with similar-sounding verbal labels (such as “cat”, “mat”, “hat”) are recalled significantly less well than pictures that have dissimilar-sounding verbal labels (such as “bell”, “shoe”, “drum”). This “phonological similarity effect” (PSE) is clear evidence that visually-presented information has been recoded into a verbal form, such that recall is affected by manipulations to the phonological properties of the to-be-remembered pictures (see Gathercole, 1998).

A number of authors have argued that it is only from approximately seven years of age onwards that typically developing children show a PSE for visually-presented material in serial recall, suggesting that before this age they do not spontaneously employ inner speech as a means of mediating short-term memory (Halliday, Hitch, Lennon, & Pettipher, 1990; Hayes & Schulze, 1977; Hitch, Halliday, Schaafstal, & Heffernan, 1991). Indeed, rather than being negatively affected by the phonological similarity of items to be recalled, children below seven years of age tend to recall items that have similar visual appearances (e.g., pen, knife, tie, all presented at the same angle of orientation) significantly less well than visually dissimilar items. This “visual similarity effect” is seen as further evidence that young children are restricted to representing items visually in short-term memory (Brown, 1977; Hayes & Shulze, 1977; Hitch, Halliday, Schaafstal, & Schraagen, 1988; Hitch, Woodin, & Baker, 1989).

An alternative way of assessing whether short-term memory (or any other aspect of cognition) is verbally mediated is to assess the effect on serial recall of *preventing* the use of inner speech during the presentation of stimuli. “Articulatory suppression” involves articulating a word or phrase repeatedly, and is thought to selectively disrupt verbal thinking (Murray, 1967), leaving visuo-spatial reasoning uninterrupted (e.g., Hyun & Luck, 2007). If an individual mediates a cognitive task verbally, then performing the task under conditions of articulatory suppression should detrimentally affect their performance, whereas it should have little impact on the performance of an individual who does not employ verbal mediation.

Several studies have shown that articulatory suppression has a substantial detrimental effect on serial recall among children from approximately six or seven years of age, but little or no impact on the serial recall of younger children (e.g., Ford & Silber, 1994; Halliday et al., 1990; Hitch & Halliday, 1983). Indeed, when inner speech is blocked by articulatory suppression, older children and adults show a pattern of serial recall that resembles the pattern observed in young children under normal conditions. Hence, articulatory suppression minimises or eliminates the PSE in older individuals (e.g., Cowan, Cartwright, Winterowd, & Sherk, 1987; Ford & Silber, 1994; Hasselhorn & Grube, 2003; Hitch et al., 1990; see also Tam, Jarrold, Baddeley, & Sabatos-DeVito, 2010), and also results in a significant visual similarity effect (Hitch et al., 1989). These findings complement those from studies assessing phonological and visual similarity effects, and support the view that short-term memory is fully verbally mediated only from approximately seven years of age onwards.

Implicit in Vygotsky's (1987) theory is the idea that the shift to fully (internalised) verbal thinking at around seven years of age is a domain-*general* one, such that multiple domains of cognition become verbally mediated at this age. This idea has been stated explicitly by several contemporary Vygotskian theorists (e.g., Fernyhough, 1996; Al-Namlah et al., 2006) and has received support from studies that have shown higher-order executive functions, such as planning and task switching, to be verbally mediated from this age onwards. For instance, two studies have shown that articulatory suppression disrupts planning abilities in typically developing children and adults (Lidstone, Meins, & Fernyhough, 2010; Wallace, Silvers, Martin, & Kenworthy, 2009; but see Phillips Wynn, Gilhooly, Della Sala, & Logie, 1999). In these studies, planning skills were assessed using the classic Tower of London task (Shallice, 1982), which consists of three coloured disks that can be arranged on three individual pegs. The aim of the task is to transform one arrangement of disks (the start state) into another arrangement (the goal state) by moving the disks between the pegs, one disk at a time. To achieve this in as few moves as possible, which is the aim of the task, requires efficient planning (e.g., Owen et al., 1990). Wallace et al. found that typically developing adolescents took significantly more moves to complete the Tower of London task under conditions of articulatory suppression than under silent conditions. Similarly, Lidstone found that 7- to 10-year old children completed significantly fewer Tower of London puzzles in the minimum number of moves when completing the task under suppression than when completing the task in silence.

Several other studies have shown that articulatory suppression negatively affects typically developing individuals' ability to switch flexibly between different cognitive activities. It is well established that switching from one task to another (e.g., subtracting one number from another on one trial and adding up two numbers on the following trial, in an alternating fashion) results in a significant increase in overall completion time, relative to undertaking the same task repeatedly (e.g., adding numbers together on successive trials) (see Monsell & Driver, 2000). The difference in completion time between task-switch and task-repeat trials is known as the "switch cost". This switch cost is significantly larger under articulatory suppression than under silent conditions (e.g., Baddeley et al., 2001; Emerson & Miyake, 2003). Thus, although articulatory suppression has only a minimal effect on performance on task-repeat trials, it has a substantial negative effect on performance on task-switch trials (Miyake, Emerson, Padilla, & Ahn, 2004).

Finally, some direct evidence for the idea that the developmental shift to verbal mediation is domain-general comes from a study by Al-Namlah et al. (2006). They found that, among a group of children with a mean age of six years, the amount of task-relevant *private* speech used during the Tower of London task was significantly associated with the size of the phonological similarity effect shown by these participants in a short-term memory task. Therefore, among typically developing children, it appears that once verbal mediation is employed for short-term memory, it is also used for higher-order planning.

Verbal mediation among individuals with autism spectrum disorder

Autism spectrum disorder is diagnosed on the basis of a set of core impairments in social engagement, communication, and behavioural flexibility (American Psychiatric Association, 2000; World Health Organisation, 1992). By definition, individuals with ASD engage in relatively little of the early communicative exchanges that Vygotsky (1934/1987) suggested were critical for the formation of verbal thinking. From a Vygotskian perspective, then, individuals with ASD would be expected to show a diminished tendency to employ inner speech as a primary means of thinking (Fernyhough, 1996; 2008). Indeed, this diminution should be apparent across multiple domains of cognition, if the shift from non-verbal to verbal mediation is a domain-general one.

Several independent facts make plausible the suggestion that inner speech use may be diminished in ASD, and that this diminution may be related to the behavioural features and cognitive deficits associated with the disorder. Firstly, individuals with ASD sometimes report a tendency toward visual thinking (or "thinking in pictures"; Grandin, 1995), and a relative or total absence of inner speech (Hurlburt, Happé, & Frith, 1994). Secondly, individuals with ASD often display the kinds of limitation in self-regulation and cognitive flexibility that are associated with diminished inner speech use in other populations (see Hill, 2004; Kenworthy, Yerys, Anthony, & Wallace, 2008). Indeed, Russell, Jarrold, and Hood

(1999) suggested that the specific profile of executive dysfunction that they argued characterised ASD might be caused by a diminished propensity to employ inner speech. They argued that individuals with ASD are reliably impaired only on those executive functioning tasks that require the maintenance in mind of novel, arbitrary information/rules. Russell et al. argue that performance on such tasks is facilitated by the use of inner speech as a tool for self-reminding about which information to follow and which information to ignore. If so, a relative lack of inner speech use by individuals with ASD could explain the deficits in executive functioning that are frequently observed among people with ASD (Hill, 2004).

Despite the strong theoretical reasons to expect a diminution of inner speech use among people with ASD (e.g., Fernyhough, 1996), controlled experimental studies have yielded an inconsistent pattern of results. Recently, Williams, Happé, and Jarrold (2008; see also Russell, Jarrold, & Henry, 1996) found that children with ASD showed a developmentally appropriate pattern of verbal mediation of short-term memory (but see Joseph, Steele, Meyer, & Tager-Flusberg, 2005). On the one hand, children with and without ASD who had a verbal mental age of seven years and above showed a large, statistically significant PSE in their serial recall of visually-presented information. On the other hand, children with and without ASD who had a verbal mental age below seven years showed no sign of a PSE, but did show a large visual similarity effect, indicating the visual mediation of short-term memory.

Winsler, Abar, Feder, Schunn, and Rubio (2007) also found that aspects of executive functioning appear to be appropriately verbally mediated among individuals with ASD. Winsler et al. assessed the amount and kind of private speech used by intellectually high-functioning children with and without ASD during tests of executive set-shifting (the Wisconsin Card Sort Task; Harris, 1990) and planning (the “Building Sticks task”; Schunn & Reder, 1998). Contrary to their expectations, Winsler et al. found children with ASD were as likely as typically developing comparison children to employ private speech during these tasks. Moreover, this private speech was both task-relevant and associated with task performance. These findings led Winsler et al. (p.1361) to conclude that “when directly examined, high-functioning children with ASD do not appear to have a deficit in the spontaneous production of relevant, potentially helpful PS [private speech] during EF [executive functioning].”

Together, the studies by Williams et al. (2008), Winsler et al. (2007), and Russell et al. (1996) suggest that verbal mediation of both short-term memory and executive functioning is typical in ASD. However, contrary to these findings, other studies have reported that articulatory suppression does *not* negatively affect the performance of individuals with ASD on measures of executive functioning or working memory, suggesting diminished verbal mediation (Holland & Low, 2010; Wallace et al., 2009; Whitehouse et al., 2006). Nevertheless, potential concerns about each of these latter studies might lead to caution over the interpretation of their results.

Two studies have explored the verbal mediation of task-switching in ASD. In Whitehouse et al. (2006), participants with ASD, as well as verbal age-matched (but not chronological age-matched) comparison participants, completed an arithmetical task-switching task, once under silent conditions and once under conditions of articulatory suppression. Whitehouse et al. report that articulatory suppression had only a minimal effect on the switching performance of children with ASD, but a significant negative effect on the switching performance of comparison participants. From this, they concluded that "the present finding that blocking inner speech use has no effect on the task-switching performance of those with autism indicates that this population does not use inner speech to complete such tasks." (p.863) However, upon re-inspection, the results turned out to be more complex than suggested by this conclusion.

In a reanalysis of Whitehouse et al.'s (2006) data, Lidstone, Fernyhough, Meins, and Whitehouse (2010) found that 60% (n = 12/20) of the ASD sample *was* substantially negatively affected by articulatory suppression, indicating that the majority of the group were employing inner speech to mediate the experimental task. The original result reported by Whitehouse et al., which indicated the ASD group were less affected by articulatory suppression than the comparison group, had been driven by only a minority of the ASD group whose task-switching performance was relatively unaffected by articulatory suppression (see Williams & Jarrold, 2010). Moreover, Lidstone et al.'s analysis highlighted that children with ASD in Whitehouse et al.'s study who were unaffected by articulatory suppression had a mean verbal mental age of only seven years and nine months (*S.D.* = 1;4). Given that children (with or without ASD) would not be expected to employ verbal mediation until their verbal mental age exceeded seven years (Williams et al., 2008), it is not necessarily atypical for a number of this developmentally young subsample to have been unaffected by articulatory suppression (e.g., Ford & Silber, 1994).

As in Whitehouse et al. (2006), Holland and Low (2010) reported that the task-switching performance of children with ASD was not significantly negatively affected by concurrent articulatory suppression. However, the groups of participants in Holland and Low's study were not closely matched for age (or, as a result, verbal IQ). Although the authors report that the difference in age between the groups was non-significant, our calculations suggest that the difference was substantial (d = 0.84). Importantly, within the ASD group, chronological age was also moderately correlated (r = -.37) with the main variable of interest, namely with the extent to which articulatory suppression negatively affected task-switching performance. As such, differences between the groups in chronological age could well have contributed to the group difference in the use of inner speech to mediate the experimental task. Moreover, Holland and Low did not present data on what proportion of the ASD group were unaffected by articulatory suppression. Therefore, as was the case with Whitehouse et al.'s results, differences between the groups in Holland and Low's study could have been driven by a small minority of the ASD group.

Two studies have explored the verbal mediation of planning in ASD. In Wallace et al. (2009), closely-matched groups of ASD and comparison participants completed four trials of a standard Tower of London task under silent conditions and four different trials under articulatory suppression. Wallace et al. reported that articulatory suppression significantly negatively affected the planning performance of comparison participants (with a small effect size; d = 0.47), but did not significantly impair the performance of ASD participants (d = 0.21). However, the interaction between the effect of articulatory suppression and diagnostic group was not significant. Consequently, the extent to which the planning performance of participants with ASD was negatively affected by articulatory suppression was *not* reliably different from the extent to which the planning performance of comparison participants was negatively affected. Indeed, according to our calculations, the difference between the groups in this respect was minimal (d = 0.29).

Holland and Low (2010) also gave participants with and without ASD a Tower of Hanoi planning task under conditions of articulatory suppression, as well as under silent conditions. Unlike Wallace et al. (2009), Holland and Low (2010) did find a significant interaction between diagnostic group and condition, which reflected the fact that typically developing comparison participants were more negatively affected by articulatory suppression than were participants with ASD. However, the Tower of Hanoi methodology employed by Holland and Low was somewhat questionable. Holland and Low employed a standard Tower task (Delis, Kaplan, & Kramer, 2001), which consists of nine trials of increasing difficulty (i.e., an increasing minimum number of moves required to complete each trial). Yet, participants completed only a single trial of the Tower task under each of the three conditions (silence, articulatory suppression, spatial tapping). In fact, from the description of the procedure provided in the paper, it appears that participants completed the *same trial* in each condition (i.e., completed the same trial on three occasions). Although the order in which each condition was undertaken was counterbalanced across participants, the fact that the same trial was completed on multiple occasions provides reason to be cautious about interpreting the processes underlying task performance.

Rationale for and details of the current study

There is a clear debate about the nature of verbal mediation in ASD. For a number of reasons, it is important that the discrepancies in results between previously-conducted studies are clarified. Perhaps most notably, if inner speech is *not* used by people with ASD as a primary means of thinking, intervention efforts could be targeted at encouraging verbal mediation with the aim of remediating aspects of the cognitive and behavioural phenotype of ASD (Williams & Jarrold, 2010). Such a strategy has proven useful for increasing cognitive flexibility among young typically developing children (Asarnow & Meichenbaum, 1979; Kray, Eber, & Karbach, 2008). However, it is far from clear that individuals are atypical in their use of verbal mediation. If individuals with ASD are typical in this respect, this would have significant consequences for theories of both typical and atypical development (Vygotsky, 1987; Fernyhough, 2008; Russell et al., 1999).

Wallace et al. (2009) suggest that one way to clarify the discrepancies between studies conducted to date is to employ a combination of tasks and techniques (used *across* various previous studies) among the same individuals. Therefore, we explored the verbal mediation of both short-term memory (Experiment 1) and executive planning (Experiment 2), assessing the effects of phonological similarity and articulatory suppression (as measures of verbal mediation) on task performance, among individuals with and without ASD. This allowed us to evaluate whether contradictory results between studies are due to (among other possibilities):

a) The differing domains of cognition assessed across previous studies: Perhaps individuals with ASD are atypical in the sense that they employ inner speech for some purposes (e.g., short-term memory), but not for other purposes (e.g., planning or task-switching). If this is the case, individuals with ASD should show different patterns of performance across Experiments 1 and 2. For example, if individuals with ASD employ inner speech for the purposes of short-term memory, but not planning, then they should show a significant PSE *and* a significant articulatory suppression effect in Experiment 1, but be unaffected by articulatory suppression in Experiment 2;

b) The relative sensitivity of different techniques to diminished inner speech use in ASD: Perhaps independent of the domain of cognition assessed, articulatory suppression is more sensitive to diminished inner speech use in ASD than are other techniques, such as similarity effects. This would explain why previous studies that have employed articulatory suppression have reported diminished inner speech use among people with ASD, whereas studies employing other techniques have found no evidence of such diminution. Hence, if inner speech use is diminished in all respects among people with ASD, but only articulatory suppression is sensitive enough to detect this, then participants with ASD should show a significant PSE in Experiment 1, but be unaffected by articulatory suppression in both Experiment 1 and Experiment 2;

c) Potential flaws in one or more of the studies: Perhaps inconsistent results between previous studies have been due to difficulties with previous study designs, rather than inherent differences in inner speech use between ASD and comparison groups. If this is the case, the sample of participants with ASD in the current study should perform similarly across both experiments. Participants with ASD may display entirely typical inner speech use and show a PSE in Experiment 1 and an articulatory suppression effect in experiments 1 and 2. Alternatively, they may show consistently diminished inner speech use and thus fail to display a PSE in Experiment 1 or an articulatory suppression effect in either Experiment 1 or Experiment 2.

Experiment 1

Method

Participants

Ethical approval for the study was obtained from City University Research Ethics Committee. Seventeen adults with ASD and 17 typically developed comparison adults took part in Experiment 1, after they had given their written, informed consent. Participants in the ASD group had received formal diagnoses of autistic disorder or Asperger's disorder, according to conventional criteria (American Psychiatric Association, 2000; World Health Organisation, 1992). All participants with ASD completed the Autism-spectrum Quotient (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), a self-report measure of ASD features, and all were administered the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000), a detailed observational assessment of ASD features. All but one comparison participant completed the AQ. Participants in the ASD group scored above the defined cut-off for ASD on both the ADOS (Total score ≥ 7; Lord et al.) and the AQ (Total score ≥ 26; Woodbury-Smith, Robinson, Wheelwright, & Baron-Cohen, 2005)[1]. The mean ADOS Total score of the ASD group was in the autism range. Participants in the comparison group scored below the defined cut-off for ASD on the AQ. No participant in either group reported any current use of psychotropic medication or illegal recreational drugs, and none reported any history of neurological or psychiatric illness, other than ASD. Using the Wechsler Adult Intelligence Scales – Third Edition UK (WAIS; Wechsler, 2000), the groups were equated for verbal, non-verbal, and full-scale IQ. The groups were also equated for chronological age. Participant characteristics are presented in Table 1.

Apparatus and stimuli

Stimuli for the serial recall task were 18 pictures similar to those used by Hitch et al. (1989) and Williams et al. (2008). Nine of the pictures had phonologically similar labels (bat, cat, hat, mat, map, rat, tap, cap), and nine control pictures had phonologically dissimilar labels (drum, shoe, fork, bell, leaf, bird, lock, fox). All items were one syllable in length and matched for word frequency as indexed by Kucera and Francis (1967) and Thorndike and Lorge (1944) counts, and for imageability and concreteness as reported in the MRC Psycholinguistic Database (Coltheart, 1981). A multivariate analysis of these four measures across the two stimulus types revealed a non-significant main effect of stimulus type using Wilks' criterion, $F(4, 10) = 0.60$, $p = .67$, confirming the adequacy of this matching. Thirteen of the 18 pictures were drawn from Snodgrass and Vanderwart's (1980) standardised set. Five of the pictures (tap, rat, cap, mat, map) were not available from Snodgrass and Vanderwart's set and so were selected from Microsoft Clipart so as to match as closely as possible the style of Snodgrass and Vanderwart's pictures. All stimuli were presented on a Dell 15 inch flat-screen monitor, using Microsoft Powerpoint.

Design and procedures

Short-term memory for the materials of each stimulus type (phonological, control) was assessed using an incremental span procedure. Items were presented in sequences that varied from two to eight pictures. There were three trials at each sequence length. Items in each trial appeared in the centre of the screen for one second. After presentation of the last item in each trial, the screen went blank and the participant was invited to recall the items in serial order. Each trial was considered to have been successfully completed if all items were recalled in correct order. If at least one of the three trials at a given sequence length was successfully completed, the participant was given another set of (three) trials at a greater sequence length. When none of the trials at a given sequence length was successfully completed, the participant moved on to the next stimulus type. The order in which each stimulus type was completed was counterbalanced across participants. Trials involving each stimulus type began with three-item sequences (i.e., three trials of three items). If none of the trials was successfully completed at this sequence length then participants were given a set of trials with two-item sequences.

Participants completed trials involving each type of stimulus under two conditions: In counterbalanced order, participants completed each stimulus type once under articulatory suppression and once under silent conditions. To illustrate, a participant might complete the phonological trials under silent conditions followed immediately by the control trials under silent conditions. Then, after a short break, they would complete a different set of phonological trials (containing the same pictures, but arranged into different sequences) under silent conditions followed immediately by a different set of control trials (containing the same pictures, but arranged in different sequences) under articulatory suppression. In the articulatory suppression condition, participants repeated either the word "Tuesday" or the word "Thursday" (counterbalanced across participants) in time to a metronome, which was set to a rate of 65 beats per minute. The metronome remained on during the silent condition, but participants did not articulate the task-irrelevant word. It is important to note that throughout both experiments reported in this study, participants from each group engaged in articulatory suppression appropriately during the suppression conditions. The experimenter was vigilant in making sure that participants articulated the task-irrelevant word in time to the metronome.

Participants were tested individually in a sound-attenuated laboratory at the university at which the research was conducted. The experimenter first showed participants each picture from the task and labelled it. This was done in order to ensure that the correct labels, which had been matched for syllable length, were being used. If, for example, a participant had consistently used an incorrect/alternative label (e.g., ‘padlock’, for the item ‘lock’) in a particular condition, then any findings would be confounded by uncontrolled “word length effects” (Baddeley, Thompson, & Buchanan, 1975). In fact, during the task participants always employed the correct terms for the pictures, without exception.

Before beginning the task, participants were given three practice trials (each involving three-item sequences) under each of the conditions (silent and suppression). Specifically, participants who completed the task under articulatory suppression first received three practice trials under silent conditions, followed immediately by three practice trials under articulatory suppression. They then completed the experimental trials with a short break between conditions. Participants who completed the silent condition first of all completed three practice trials under silent conditions before completing the experimental trials under silent conditions. Then, after a short break, they completed three practice trials under articulatory suppression, before completing the experimental trials under suppression.

Scoring

Participants’ recall performance was determined using a “partial credit scoring” method, which is considered the gold-standard way to score memory span (Conway et al., 2005). According to this method, participants received a score of one for every trial in which all items were correctly recalled in serial order, plus a proportional score for each unsuccessful trial. This proportional score corresponded to the proportion of items within each trial that were recalled in the correct position. Hence, if a participant recalled two out of four items (on a four-item trial), their score for that trial would be .50.

Two scores were employed as measures of inner speech use. First, the size of the *phonological similarity effect* was determined by subtracting recall performance on phonological trials completed under silent conditions from recall performance on control trials completed under silent conditions. The more positive the resulting value, the more one can assume that inner speech was relied upon to complete the task. Second, the size of the *articulatory suppression effect* was determined by subtracting recall performance on control trials completed under silent conditions from recall performance on control trials completed under articulatory suppression. Again, the more positive the resulting value, the greater the evidence that inner speech was relied upon to complete the task.

As argued above, in addition to analysing group means, we also believe it is important to explore individual data. Therefore, we created two categorical variables that corresponded to the PSE and articulatory suppression effect, respectively. Categorically, participants were deemed to have shown a PSE if they recalled ≥ 1 trial more on control trials than on phonological trials. Likewise, participants were deemed to have shown an articulatory suppression effect if they recalled ≥ 1 trial more on control trials in the silent condition than on control trials in the articulatory suppression condition.

Results

Table 2 shows the mean number of trials correctly recalled by ASD and comparison participants in each condition (suppression/silent), by stimulus type (phonological/control). A mixed ANOVA was conducted on these data, with Condition and Stimulus type as within-participant variables, and Group as the between-participants variable. There was a significant main effect of Condition, $F(1, 32) = 40.00$, $p < .001$, and a significant main effect of Stimulus type, $F(1, 32) = 51.39$, $p < .001$. However, these main effects were qualified by a significant interaction between Condition and Stimulus type, $F(1, 32) = 24.67$, $p < .001$.

To break down this interaction, a series of within-participant *t*-tests exploring the recall of trials involving each stimulus type in each condition was conducted. In the silent condition, phonologically similar stimuli were recalled significantly less well than control stimuli, indicating a clear phonological similarity effect, with a large effect size, $t(33)= -6.65$, $p < .001$, $d = -1.39$. In contrast, in the articulatory suppression condition, phonologically similar stimuli were recalled non-significantly less well than control stimuli, with only a small effect size, $t(33)= -1.98$, $p = .06$, $d = -0.19$. Hence, as predicted, a significant phonological similarity effect was apparent in the silent condition, but not the articulatory suppression condition. Also, recall of control stimuli in the articulatory suppression condition was significantly poorer than recall of control stimuli in the silent condition, indicating a clear articulatory suppression effect, with a large effect size, $t(33)= 6.36$, $p < .001$, $d = -1.34$. In contrast, recall of phonologically similar stimuli in the suppression condition was only marginally significantly poorer than recall of phonologically similar stimuli in the silent condition, with a small effect size, $t(33)= 2.02$, $p = .05$, $d = -0.25$. Hence, as predicted, articulatory suppression had a substantial negative effect on the recall of control stimuli, but only a marginal effect on the recall of phonologically similar stimuli.

There was no significant main effect of Group, $F(1, 32) = 0.87$, $p = .36$, and no significant interaction between Group and Condition, $F(1, 32) = 0.71$, $p = .41$, or between Group and Stimulus type, $F(1, 32) = 0.58$, $p = .45$. The three-way interaction between Group, Condition, and Stimulus type was also non-significant, $F(1, 32) = 0.04$, $p = .85$. Therefore, participants with ASD were similar to comparison participants in terms of both overall levels and patterns of performance (see Figure 1).

Categorically, 15/17 (88%) participants with ASD and 16/17 (94%) comparison participants showed a PSE. In this respect, the groups were not different, $\chi^2 = 0.37$, Fisher's Exact $p > .99$, $\varphi = .10$. Similarly, 13/17 (76%) participants with ASD and 14/17 (82%) comparison participants showed an articulatory suppression effect, $\chi^2 = 0.18$, Fisher's Exact $p > .99$, $\varphi = .07$.

Associations between inner speech use and ASD features

A series of correlation analyses was conducted to explore the relation between the key experimental measures of verbal mediation (size of PSE and size of articulatory suppression effect), as well as the relations between each of these measures, respectively, and ASD features (as measured by the ADOS and AQ). Firstly, when analysing the continuous data, the size of the PSE was significantly associated with the size of the articulatory suppression effect among both ASD participants, $r_s = .88$, $p < .001$, and comparison participants, $r_s = .74$, $p = .001$. When analysing the categorical data, there was a significant association between displaying a PSE and displaying an articulatory suppression effect among participants from both diagnostic groups, $\chi^2 = 4.27$, Fisher's Exact $p = .04$, $\varphi = .36$.

The ADOS has a "Total score", which is a combination of scores from two core diagnostic subscales, the "Reciprocal Social Interaction" subscale and the "Communication" subscale. To adjust for multiple comparisons in analyses involving the ADOS, a Bonferroni corrected alpha level of <.017 was applied. Among participants with ASD, neither the size of the PSE nor the size of the articulatory suppression effect was significantly associated with the ADOS Total score, or either of the core ADOS subscale scores (all r_ss < .29, all ps > .27).

The AQ has a Total score, which is derived from scores on five subscales: "Social Skill", "Attention Switching", "Attention to Detail", "Communication", "Imagination". To adjust for multiple comparisons in analyses involving the AQ, a Bonferroni corrected alpha level of <.008 was applied. The size of the PSE was not significantly associated with the AQ Total score, or any of the five subscale scores among participants with ASD (all r_ss < .40, all *p*s > .11), or among comparison participants (all r_ss < .32, all *p*s > .23). Similarly, the size of the articulatory suppression effect was not significantly associated with the AQ Total score, or any of the five AQ subscale scores among participants with ASD (all r_ss < .55, all *p*s > .02), or among comparison participants (all r_ss < .45, all *p*s > .08).

Discussion

The results of Experiment 1 were clear. Participants from each group showed a substantial PSE in serial recall, indicating that (visually presented) stimuli were spontaneously recoded, and then presumably rehearsed, prior to recall. This result replicates that of Williams et al. (2008) and arguably confirms their suggestion that individuals with ASD are typical in employing inner speech as a means of retaining information in short-term memory. Nonetheless, as highlighted above, it could have been that the PSE was insensitive to diminished inner speech use in ASD and that Williams et al.'s failure to find differences between their groups of participants was merely an artefact of this insensitivity (Wallace at al., 2009). However, in the current study, recall performance among both participants with ASD and comparison participants was also substantially negatively affected by articulatory suppression. Indeed, among each group of participants, the degree to which phonological similarity of items negatively affected recall performance was highly correlated with the degree to which articulatory suppression negatively affected performance. This suggests that both measures were assessing a common underlying process in each group of participants, namely the degree to which inner speech was relied upon to mediate the experimental task. Also, at the individual level almost 90% of participants with ASD showed a PSE and almost 80% showed an articulatory suppression effect. Together, these results provide convincing evidence that short-term memory for nameable visually-presented information is verbally mediated among the majority of people with ASD who have a verbal mental age of seven years or above (cf. Williams et al., 2008). What remains unclear, however, is whether people with ASD rely less than people without ASD on inner speech use for purposes other than retaining information in short-term memory. Experiment 2 explored whether the same participants with ASD also employ inner speech for the purpose of planning.

Experiment 2

Participants

Fifteen participants with ASD and 16 comparison participants took part in Experiment 2. All of these participants also took part in Experiment 1. Two participants from the ASD group and one comparison participant elected not to take part in Experiment 2. The groups were matched for age, VIQ, PIQ, and FSIQ (all *t*s < 1.25, all *p*s > 0.22, all *d*s < 0.45). The mean AQ score of the ASD group (*M* = 34.53, *S.D.* = 7.24) was significantly higher than that of the comparison group (*M* = 12.13, *S.D.* = 5.86), $t = 9.50$, $p < .001$, $d = 3.40$.

Apparatus and stimuli

Participants completed 18 computerised Tower of London puzzles, each involving three pegs, and five coloured disks of different sizes (see Figure 2). Each puzzle was presented on a 14-inch Dell laptop screen. The goal state was visible throughout each trial at the top of the screen. Directly underneath the goal state was the puzzle for participants to complete, which always began in the appropriate starting state.

The puzzles were selected from those listed in Appendix B of Ward and Allport (1997, p.77). Puzzles were divided into two sets, each consisting of nine puzzles. Across sets, the puzzles were equated for difficulty in terms of the minimum number of moves required to solve each (i.e., reach the goal state from the start state). In each set, two problems required a minimum of five moves to reach a solution, two required a minimum of seven moves, two required a minimum of nine moves, and one puzzle in each set required a minimum of 10, 11, and 13 moves, respectively.

Ward and Allport identified two further factors that influence the relative difficulty of ToL problems; the number of "subgoal moves" required and the number of "subgoal chunks" required. A subgoal move is defined by Ward and Allport as "a move that is essential to the optimum solution, but which does not place a disk into its goal position" (p.56). A subgoal chunk is defined as "a consecutive series of subgoal moves that transfer disks to and from the same pegs" (p.57). Ward and Allport found that, among typical adults, as each of these factors increased so did the number of errors (i.e., non-optimal moves), indicating an increasing load on planning resources. As such, in the current study, puzzles in each set were also matched for number of subgoals (ranging from zero to five, per puzzle) and the number of subgoal chunks (ranging from zero to four per puzzle).

Design and procedures

Participants completed one set of puzzles under silent conditions and the other set of puzzles under concurrent articulatory suppression. The order in which the conditions (suppression and silent) were completed, as well as the order in which sets of puzzles were presented, was counterbalanced across participants. In the articulatory suppression condition, participants repeated either the word "Tuesday" or the word "Thursday" (counterbalanced across participants) in time to a metronome, which was set to a rate of 65 beats per minute.

Before beginning the experimental trials, participants were given three practice trials (involving two, three, and four move sequences, respectively) under each of the conditions (silent and suppression). In the same manner as in Experiment 1, those participants who first undertook the articulatory suppression condition completed three practice trials in silence, followed by three practice trials under suppression, before beginning the experimental trials. Participants who first undertook the silent condition completed three practice trials in silence, before completing the experimental trials under silent conditions. Then, after a short break, they completed three practice trials under articulatory suppression, before completing the experimental trials under suppression.

Participants were introduced to the task by the experimenter, who explained that the aim was to "make the puzzle at the bottom of the screen (start state) look exactly like the puzzle at the top of the screen (goal state)." On a single trial, the experimenter demonstrated how the disks could be moved from peg to peg, and explained how any disk could go on top of any other disk. All participants understood the nature of the task. The experimenter explained, further, that the "aim was to complete the puzzle in as few moves as possible. So, you'll need to plan how to move the disks before you start."

Scoring

An articulatory suppression effect index was created by subtracting the total number of moves taken to complete the puzzles in the silent condition from the total number of moves taken to complete the puzzles in the articulatory suppression condition. The more positive the resulting value, the more it was assumed that inner speech was relied upon to complete the task. Categorically, participants were deemed to have shown an articulatory suppression effect if they if they took ≥ 1 more move to complete puzzles in the articulatory suppression condition than in the silent condition.

Results

Participants with ASD took an average of 84.87 (S.D. = 5.04) moves to complete all nine Tower puzzles in the silent condition and an average of 84.47 (S.D. = 5.79) moves to complete all nine in the articulatory suppression condition. Comparison participants took an average of 83.44 (S.D. = 5.93) moves to complete the puzzles in the silent condition and an average of 89.25 (S.D. = 6.52) moves to complete the puzzles in the articulatory suppression condition. A mixed ANOVA was conducted on these data, with Condition (suppression/silent) as the within-participant variable and Group as the between-participants variable. The main effect of Group was non-significant, $F(1, 29) = 1.03$, $p = .32$. There was a significant main effect of Condition, $F(1, 29) = 4.32$, $p = .05$. However, this was qualified by a significant interaction between Condition and Group, $F(1, 29) = 5.69$, $p = .02$. To break down this interaction, within- and between-participant *t*-tests were conducted exploring performance in each condition. Whereas participants with ASD performed comparably in each condition (i.e., were not negatively affected by articulatory suppression), $t(14) = 0.20$, $p = .85$, $d = 0.07$, comparison participants performed significantly less well in the articulatory suppression condition than in the silent condition, $t(15) = 3.46$, $p = .003$, $d = -0.93$. Indeed, whereas participants with ASD performed non-significantly *less* well than comparison participants in the silent condition, $t(29) = 0.72$, $p = .48$, $d = -0.26$, they performed significantly *better* than comparison participants in the articulatory suppression condition, $t(29) = 2.15$, $p = .04$, $d = 0.78$.

Categorically, 6/15 (40%) participants with ASD and 14/16 (88%) comparison participants showed an articulatory suppression effect. In this respect, the groups were significantly different, $\chi^2 = 7.63$, $p = .006$, $\varphi = .50$.[2]

Associations between inner speech use and ASD features

A series of correlation analyses was conducted to explore the relations between the key experimental measure of verbal mediation (size of articulatory suppression effect) and ASD features. As in Experiment 1, a Bonferroni corrected alpha level of <.017 was applied in analyses involving the ADOS. Among participants with ASD, the size of the articulatory suppression effect was not significantly associated with the ADOS Total Score, $r_s = -.41$, $p = .13$, or with the ADOS Reciprocal Social Interaction subscale score, $r_s = -.07$, $p = .81$.

However, the size of the articulatory suppression effect was strongly and significantly associated with the ADOS Communication subscale score $r_s = .72$, $p = .003$.

As in Experiment 1, a Bonferroni corrected alpha level of <.008 was applied in analyses involving the AQ. Among participants with ASD, the size of the articulatory suppression effect was significantly associated with AQ Communication subscale score, $r_s = .76$, $p = .001$, only (all other r_ss < .56, all other *p*s > .03).[3] It is important to note that the correlations between the size of articulatory suppression effect, and ADOS Communication score and AQ Communication score, respectively, were not merely a by-product of verbal intelligence, given that verbal IQ was not positively or significantly associated with any of these variables (all *r*s < -.31, all *p*s > .23).

Next, through a series of linear regression analyses, we assessed the extent to which variance in the articulatory suppression effect was explained by a common factor underlying both the ADOS Communication subscale score and the AQ Communication subscale score. Together, the two scores explained 69.3% of the variance in the size of the articulatory suppression effect. The ADOS Communication subscale score uniquely accounted for 12.6% of the variance and the AQ Communication subscale score uniquely accounted for 27.3% of the variance. Thus, 29.4% of the variance in the size of the suppression effect was explained by an underlying factor shared by the two subscale scores. In other words, a fundamental aspect of communication ability, assessed by *both* measures, was driving the significant correlations between the size of the suppression effect and each of the subscale scores, respectively.

Post-hoc analysis of successful planning among individuals with ASD

The finding that participants with ASD did not perform significantly less well than comparison participants in the silent condition of the Tower task was not unexpected, given that several studies of planning abilities in ASD have reported null results when using computerised versions of this planning task (Goldberg et al. 2005; Happé, Booth, Charlton, & Hughes, 2006; Just Cherkassky, Keller, Kana, & Minshew, 2007; Ozonoff et al. 2004). However, what needs to be explained is *how* individuals with ASD are performing well on the task in the current study, given that they were not apparently employing inner speech to mediate their planning. According to Motton and colleagues (e.g., Caron, Mottron, Bethiaume, & Motton, 2006), among others (e.g., Plaisted, O'Riordan, & Baron-Cohen, 1998), visuo-spatial abilities tend to be enhanced among individuals with ASD and employed to solve tasks that might be solved by other means among typically developing individuals. In a recent review, Mottron, Dawson, Soulières, Hubert & Burack (2006, p.39) concluded that,

> …perception plays a different and superior role in autistic cognition. Recent studies in the visual and auditory modalities indicate a skewing of brain activation toward primary and early associative areas in autistics in most tasks involving higher-order or socially relevant information…

Of the twelve subtests that comprise the WAIS, the Block Design subtest is considered to be a unique measure of visuo-spatial abilities (e.g., Caron et al., 2006). Therefore, in order to examine whether visuo-spatial abilities were uniquely associated with (and arguably underlie) planning performance among participants with ASD, we conducted correlation analyses exploring the relation between performance on the Block Design subtest of the WAIS and performance in the silent condition of the Tower task. It is important to stress that, although these analyses were *post-hoc*, they were the only analyses we conducted and they were based on the specific hypothesis that planning performance in ASD is uniquely underpinned by perceptual abilities, whereas it is uniquely underpinned by inner speech use among comparison participants. In line with this hypothesis, performance on the Block Design subtest was highly and significantly associated with planning performance among individuals with ASD, $r_s = .64$, $p = .01$. In contrast, the association was minimal among comparison participants, $r_s = .03$, $p = .92$.

Association between inner speech use in Experiment 1 and inner speech use in Experiment 2

To explore whether the use of inner speech to mediate short-term memory was associated with the use of inner speech to mediate planning, analyses were conducted to assess the relation between the PSE and the articulatory suppression effect, respectively, from Experiment 1 with the articulatory suppression effect from Experiment 2. Analysis of the continuous data revealed that among neither group of participants was there a significant association between the size of the PSE in Experiment 1 and the size of the articulatory suppression effect in Experiment 2, or between the size of the articulatory suppression effect in Experiment 1 and the size of the articulatory suppression effect in Experiment 2 (all rs $< -.23$, all ps $> .41$).

Analysis of the categorical data revealed an important difference between the diagnostic groups in patterns of performance *across* experiments. Among comparison participants, 13/16 (81%) showed a categorical PSE in Experiment 1 *and* a categorical articulatory suppression effect in Experiment 2. Likewise, 12/16 (75%) of comparison participants showed a categorical articulatory suppression effect in Experiment 1 *and* a categorical articulatory suppression effect in Experiment 2. However, among participants with ASD, the pattern of performance across experiments was quite different. Only 5/15 (33%) participants with ASD showed a categorical PSE in Experiment 1 *and* a categorical articulatory suppression effect in Experiment 2. In contrast, 8/15 (53%) participants with ASD showed a categorical PSE in Experiment 1, but *not* a categorical articulatory suppression effect in Experiment 2. This compared to only 1/15 (0.07%) of participants with ASD who showed the opposite pattern of performance. Therefore, participants with ASD were significantly more likely to use inner speech to mediate their short-term memory, but not their planning than vice versa, McNemar's $p = .04$. A similar result was observed when comparing the effects of articulatory suppression across experiments. Only 4/15 (28%) of participants with ASD showed a categorical articulatory suppression effect in Experiment 1 *and* a categorical articulatory suppression effect in Experiment 2. Instead, 7/15 (47%) showed a categorical articulatory suppression effect in Experiment 1, but not a categorical articulatory suppression effect in Experiment 2. This compared to only 2/15 (13%) participants who showed the opposite pattern, McNemar's one-tailed $p = .09$.

Discussion

The results of Experiment 2 were clear; preventing inner speech use by imposing articulatory suppression had a significant detrimental effect on the planning performance of comparison participants (d = -0.93). In contrast, preventing inner speech use among participants with ASD had next to no effect on their planning performance (d = 0.07). At the individual level, only just over one third of participants with ASD were at all negatively affected by the imposition of articulatory suppression, whereas almost 90% of comparison participants were so affected. These results suggest that individuals with ASD rely significantly less than comparison participants on inner speech to mediate their planning. *Post-hoc* analyses provided some evidence that, instead of using inner speech to mediate their planning, individuals with ASD relied on their visuo-spatial skills to mediate the Tower task.

Perhaps most strikingly, the degree to which articulatory suppression negatively affected Tower of London performance among ASD participants was highly and significantly correlated with the severity of communication difficulties experienced by these individuals. In other words, as the severity of communication difficulties increased (as established either by detailed observation, using the ADOS, or self-report, using the AQ), inner speech use for planning decreased.

General discussion

The idea that language/speech plays a significant role in thinking is increasingly (although not universally) accepted by cognitive scientists and psychologists (e.g., Carruthers, 2002). Moreover, according to Vygotsky (1987), verbal thinking has its origins in interpersonal communication with others early in life. Together, these two ideas have understandably led to the idea that a failure of verbal thinking may be implicated in ASD – arguably the prototypical disorder of social-communication, which also involves diminished higher-order cognition (e.g., Fernyhough, 1996). Empirical research on verbal thinking in ASD had produced mixed results and we raised concerns about the methodological approaches taken in those studies that claimed to have observed diminished verbal mediation in ASD. The results of the current study arguably provide a clearer picture not only of the nature of verbal thinking among people with ASD, but also of the way verbal thinking typically develops.

Broadly, the results of this study support the idea outlined above that individuals with ASD are atypical in the sense that they employ inner speech for the purpose of recoding visually-presented information into a verbal code in order to retain it in short-term memory, but do not employ inner speech to assist their planning. The findings that participants with ASD showed a clear PSE in their serial recall of visually-presented material, and that articulatory suppression severely disrupted their recall performance, provides strong support for the idea that verbal recoding of visual information is common among the majority of people with this disorder (cf. Williams et al., 2008; Williams & Jarrold, 2010). Arguably, however, the current study is the first to demonstrate convincingly that an aspect of executive functioning, namely planning, is *not* verbally mediated among the majority of people with ASD. In the current study, planning performance was not detrimentally affected by articulatory suppression among the majority of participants with ASD, unlike among comparison participants, the majority of whom were severely negatively affected. Instead, the planning performance of participants with ASD was uniquely associated with visuo-spatial processing abilities, as measured by the Block Design subtest of the WAIS. Although caution is certainly warranted when interpreting this latter result (given that the analysis that revealed this finding was conducted *post-hoc*, as well as given difficulties with inferring causation from correlation), this provides some evidence in support of Mottron et al.'s (2006, p.39) claim that "perception plays a different and superior role in autistic cognition". Specifically, this result suggests that individuals with ASD rely on visuo-spatial abilities, rather than inner speech, to mediate their planning.

One striking implication of the current findings is that the *mechanism* underpinning inner speech use is intact among people with ASD, but that fundamentally different *forms* of inner speech are involved in mediating different cognitive domains – and, critically, that only one of these forms is diminished among individuals with ASD. Following Fernyhough (1996, 2008), Williams et al. (2008, p.57) distinguished between inner speech that is "dialogic" and inner speech that is "monologic", and questioned whether individuals with ASD showed a diminution of the former kind only. As Fernyhough (2008, p.233) highlights, "the verbal thinking upon which we can sometimes introspect often appears to us as a kind of dialogue between distinct perspectives on reality." Therefore, dialogic inner speech involves a kind of "conversation" between different aspects of self/perspectives held by self and is an ideal medium for accommodating multiple, alternative perspectives upon a topic of thought. It is this ability to hold in mind and move flexibly between different perspectives on a situation that arguably facilitates efficient problem-solving in situations where one might otherwise become "stuck-in-set". This form of inner speech could clearly maximise planning efficiency on the Tower of London task by allowing one to mentally consider alternative ways of moving from the start state to the goal state, and then act according to the best mental model. However, we suggest (following Fernyhough's reading of Vygotsky) that this form of inner speech use may have inherently social origins and that without adequate experiences of communicating with others this kind of inner speech will not develop typically. The message from Vygotskian theory is clear: individuals who are poor at *conversing* with others will be poor at *conversing* with self. This would explain both why the majority of participants with ASD were unaffected by the imposition of articulatory suppression during the Tower of London task in Experiment 2, and also why the extent to which they were affected by suppression was associated closely with the severity of their communication impairments.

In contrast to dialogic inner speech, monologic inner speech involves merely a commentary by self about a particular state of affairs. This form of inner speech might be described as "for oneself", unlike dialogic inner speech which is "to oneself". The development of this kind of inner speech is far from trivial and it could have considerable benefits for cognition. For example, rehearsing novel verbal information may facilitate the acquisition of long-term knowledge by preventing its loss from short-term memory. However, this kind of verbal labelling and subvocal rehearsal is clearly not "conversational" in the same way that dialogic inner speech is. Arguably, therefore, the ability to engage in this kind of inner speech does not depend on experience of social-communicative exchanges with others. This would explain why the serial recall performance of participants with ASD *was* negatively affected by articulatory suppression and phonological similarity in Experiment 1, and also why the size of these effects was *not* significantly associated with communication skills among these participants. Moreover, the idea that only dialogic inner speech is diminished in ASD would make sense of the finding that participants with ASD in the current study used inner speech *inconsistently* across experiments. For example, participants with ASD were significantly more likely to employ inner speech in Experiment 1 only than they were to employ inner speech in Experiment 2 only. In contrast, the vast majority of comparison participants employed inner speech across both experiments. One interpretation of this is that participants with ASD are restricted to employing monologic inner speech, whereas comparison participants can engage in both monologic and dialogic forms of inner speech.

The current findings have other important implications for our understanding of the typical development and use of verbal mediation. Firstly, the evidence from ASD does not support the Vygotskian hypothesis that the shift from visual to verbal mediation is domain-general. Rather, the evidence from ASD suggests that it is possible for inner speech to be used quite typically to mediate some domains of cognition, but not other domains. This suggests that the apparent domain-generality of inner speech use among typically developing individuals (e.g., Al-Namlah et al., 2006) may only be superficial. Secondly, these results suggest that there is a critical distinction between possessing good structural language and *using* this for the purpose of structuring cognition. In the current study, participants with ASD were verbally able, but did not use inner speech to support their planning. Conversely, there is recent evidence that children with specific language impairment (SLI), who by definition have impaired structural language but comparatively unimpaired communication skills, do employ inner speech to mediate their planning (Lidstone, Fernyhough, & Meins, 2010).

The implications of the current study (outlined directly above) could be assessed in a number of ways. First, future studies should explore directly the quality of inner speech used by individuals with and without ASD to mediate different aspects of cognition. This might be done, in the first instance, via self-report (although self-reported use of inner speech by individuals with ASD may not wholly accurate; Williams et al., 2008). We predict that only dialogic forms of inner speech will be associated with communication skills. Related to this, inner speech use could be further explored among participants with language impairments, contrasting those participants in whom language impairment is primarily structural (as in SLI) with those participants in whom impairment is primarily pragmatic (as in pragmatic language impairment; Bishop, 1989). We predict that only among children with pragmatic language impairment will verbal mediation be diminished. Specifically, children with pragmatic language impairment should resemble individuals with ASD in showing diminished dialogic inner speech only.

Finally, the current results may be used to inform teaching and intervention strategies for children with ASD. First, the finding that inner speech (even if only monologic inner speech) can be employed by individuals with ASD to mediate short-term memory has implications for teaching strategies. For example, as Eley (2008) highlights, many UK-based specialist schools for children with ASD use visual timetables to support children with ASD. However, given that verbal rehearsal provides a more efficient means of scaffolding short-term memory (and, hence, long-term learning) than does visual imagery, and because individuals with ASD (who have a verbal mental age over 7 years) are capable of verbal rehearsal, it may be more productive to encourage verbal learning of timetables among these children. Second, the fact that the mechanism underlying at least some aspects of inner speech is intact among individuals with ASD leads us to wonder whether dialogic forms of inner speech might be encouraged as part of intervention efforts. Among young typically developing children, efforts to encourage monologic forms of inner speech have been somewhat successful, significantly improving children's performance on a variety of cognitive tasks (e.g., Asarnow & Meichenbaum, 1979; Kray et al., 2008). However, there is some (arguably justified) scepticism that efforts to train dialogic forms of inner speech have any meaningful long-term benefits for cognition among typically developing children (see Diaz & Berk, 1995). Nonetheless, no such training efforts have been targeted at children with ASD and we believe that there may be some value to conducting studies to explore this issue further.

What is clear from the current results is that there is not a blanket failure to employ verbal mediation among people with ASD. In certain domains of cognition, at least, there is not even a tendency for individuals with ASD to employ visual rather than verbal mediation, as some have suggested (Kunda & Goel, 2010). The short-term memory task employed in the current study (just as in the study by Williams et al., 2008) was equally amenable to visual and verbal solutions, yet participants with ASD consistently mediated the task verbally. We suggest that the likelihood of individuals with ASD employing inner speech to mediate a given cognitive task depends on the *kind* of verbal mediation that will support performance. Only in those circumstances in which truly dialogic inner speech is important for task success would we predict differences between individuals with and without ASD in underlying meditational strategies. Equally, we suggest that to explain these hypothesised differences in strategy among people with ASD will require a truly developmental perspective that explains not only the nature of differences, but also the ontogenetic origins of these differences.

References

Al-Namlah, A. S., Fernyhough, C., & Meins, E. (2006). Sociocultural influences on the development of verbal mediation: Private speech and phonological recoding in Saudi Arabian and British samples. *Developmental Psychology, 42*(1), 117-131.

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

Asarnow, J. R., & Meichenbaum, D. (1979). Verbal rehearsal and serial recall - mediational training of kindergarten-children. *Child Development, 50*(4), 1173-1177.

Austin, E. J. (2005). Personality correlates of the broader autism phenotype as assessed by the autism spectrum quotient (AQ). *Personality and Individual Differences, 38*(2), 451-460.

Baddeley, A.D. & Hitch, G.J. (1974). Working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory. Vol. VIII.*, pp. 47-90, New York: Academic Press.

Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of Experimental Psychology-General, 130*(4), 641-657.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and structure of short-term-memory. *Journal of Verbal Learning and Verbal Behavior, 14*(6), 575-589.

Baldo, J. V., Dronkers, N. F., Wilkins, D., Ludy, C., Raskin, P., & Kim, J. Y. (2005). Is problem solving dependent on language? *Brain and Language, 92*(3), 240-250.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*(6), 603-603.

Bishop, D.V.M. (1989). Autism, Asperger's syndrome, and semantic-pragmatic disorder: Where are the boundaries? *International Journal of Language and Communication Disorders, 24*, 107-121.

Brown, R. M. (1977). Examination of visual and verbal coding processes in pre-school children. *Child Development, 48*(1), 38-45.

Caron, M. J., Mottron, L., Berthiaume, C., & Dawson, M. (2006). Cognitive mechanisms, specificity, and neural underpinnings of visuospatial peaks in autism. *Brain, 129*, 1789–1802.

Carruthers, P. (2002). The cognitive functions of language. *Behavioral and Brain Sciences, 25*(6), 657-674.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology, 33*(NOV), 497-505.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769-786.

Cowan, N., Cartwright, C., Winterowd, C., & Sherk, M. (1987). An adult model of pre-school children's speech memory. *Memory & Cognition, 15*(6), 511-517.

Delis, D. C., Kaplan, E., & Kramer, J. (2001). *Delis Kaplan Executive Function System*. San Antonio, TX: The Psychological Corporation.

Diaz, R.M., & Berk, L.E. (1995). A Vygotskian critique of self-instructional training. *Development and Psychopathology, 7*, 369-392.

Dunbar, K., & Sussman, D. (1995). Toward a cognitive account of frontal lobe function: Simulating frontal lobe deficits in normal subjects. *Structure and Functions of the Human Prefrontal Cortex, 769*, 289-304.

Eley, T. Work in progress: Toward DSM-V. *Journal of Child Psychology and Psychiatry, 49,* 1-2.

Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language, 48*(1), 148-168.

Fernyhough, C. (1996). The dialogic mind: A dialogic approach to the higher mental functions. *New Ideas in Psychology, 14*(1), 47-62.

Fernyhough, C. (2008). Getting Vygotskian about theory of mind: Mediation, dialogue, and the development of social understanding. *Developmental Review, 28*(2), 225-262.

Fernyhough, C., & Fradley, E. (2005). Private speech on an executive task: Relations with task difficulty and task performance. *Cognitive Development, 20*(1), 103-120.

Ford, S., & Silber, K. P. (1994). Working-memory in children - a developmental-approach to the phonological coding of pictorial material. *British Journal of Developmental Psychology, 12*, 165-175.

Frawley, W. (2002). Inner speech and the meeting of the minds. *Behavioral and Brain Sciences, 25*(6), 686-687

Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry, 39*(1), 3-27.

Grandin, T. 1995. Thinking in Pictures. Vintage Press, New York: NY.

Goldberg, M. C., Mostofsky, S. H., Cutting, L. E., Mahone, E. M., Astor, B. C., et al. (2005). Subtle executive impairments in children with autism and children with ADHD. *Journal of Autism and Developmental Disorders, 35*, 379–293.

Gruber, O., & Goschke, T. (2004). Executive control emerging from dynamic interactions between brain systems mediating language, working memory and attentional processes. *Acta Psychologica, 115*(2-3), 105-121.

Halliday, M.S., Hitch, G.J., Lennon, B. & Pettipher, C. (1990). Verbal short-term memory in children: The role of the articulatory loop. *European Journal of Cognitive Psychology, 2*, 23-38.

Happé, F., Booth, R., Charlton, R., & Hughes, C. (2006). Executive function deficits in autism spectrum disorders and attention deficit hyperactivity disorder: Examining profiles across domains and ages. *Brain and Cognition*, *61*, 25–39.

Harris, M. E. (1990). *Wisconsin Card Sorting Test: Computer Version*. Odessa, FL: Psychological Assessment Resources.

Hasselhorn, M., & Grube, D. (2003). The phonological similarity effect on memory span in children: Does it depend on age, speech rate, and articulatory suppression? *International Journal of Behavioral Development, 27*(2), 145-152.

Hayes, D. S., & Schulze, S. A. (1977). Visual encoding in pre-schoolers' serial retention. *Child Development, 48*(3), 1066-1070.

Hill, E. L. (2004). Evaluating the theory of executive dysfunction in autism. *Developmental Review, 24*(2), 189-233.

Hitch, G. J., & Halliday, M. S. (1983). Working memory in children. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 302*(1110), 325-340.

Hitch, G. J., Halliday, M. S., Schaafstal, A. M., & Heffernan, T. M. (1991). Speech, inner speech, and the development of short-term-memory - effects of picture-labeling on recall. *Journal of Experimental Child Psychology, 51*(2), 220-234.

Hitch, G. J., Halliday, S., Schaafstal, A. M., & Schraagen, J. M. C. (1988). Visual working memory in young-children. *Memory & Cognition, 16*(2), 120-132.

Hitch, G. J., Woodin, M. E., & Baker, S. (1989). Visual and phonological components of working memory in children. *Memory & Cognition, 17*(2), 175-185.

Holland, L., & Low, J. (2010). Do children with autism use inner speech and visuo-spatial resources for the service of executive control? Evidence from suppression in dual tasks. *British Journal of Developmental Psychology, 28*(2), 369-391.

Hurlburt, R. T., Happe, F., & Frith, U. (1994). Sampling the form of inner experience in 3 adults with Asperger syndrome. *Psychological Medicine, 24*(2), 385-395.

Hurst, R. M., Mitchell, J. T., Kimbrel, N. A., Kwapil, T. K., & Nelson-Gray, R. O. (2007). Examination of the reliability and factor structure of the autism spectrum quotient (AQ) in a non-clinical sample. *Personality and Individual Differences, 43*, 1938-1949.

Hyun, J.S., & Luck, S.L. (2007). Visual working memory as the substrate for mental rotation. *Psychonomic Bulletin & Review, 14*(1), 154-158.

Joseph, R.M., Steele, S.D., Meyer, E. & Tager-Flusberg, H. (2005) 'Self-ordered pointing in children with autism: Failure to use verbal mediation in the service of working memory?' *Neuropsychologia* 43, 1400–1411.

Just, M. A., Cherkassky, V. L., Keller, T. A., Kana, R. K., & Minshew, N. J. (2007). Functional and anatomical cortical underconnectivity in autism: Evidence from an fMRI study of an executive function task and corpus callosum morphometry. *Cerebral Cortex, 17*, 951–961.

Kray, J., Eber, J., & Karbach, J. (2008). Verbal self-instructions in task switching: A compensatory tool for action-control deficits in childhood and old age? *Developmental Science, 11*(2), 223-236.

Kenworthy, L., Yerys, B. E., Anthony, L. G., & Wallace, G. L. (2008). Understanding executive control in autism spectrum disorders in the lab and in the real world. *Neuropsychology Review, 18*(4), 320-338.

Kucera, H. and Francis, W. N. (1967). *Computational Analysis of Present Day American English*. Providence: Brown University Press.

Kunda, M., & Goel, A.K. (2010). Thinking in pictures as a cognitive account of autism. *Journal of Autism and Developmental Disorders, in press (available online).*

Lidstone, J. S. M., Fernyhough, C., & Meins, E. (2010). Verbal mediation of cognition in children with specific language impairment. *Development and Psychopathology*, under revised review.

Lidstone, J. S. M., Meins, E., & Fernyhough, C. (2010). The roles of private speech and inner speech in planning during middle childhood: Evidence from a dual task paradigm. *Journal of Experimental Child Psychology, 107*(4), 438-51.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*(3), 205-223.

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. C. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica, 115*(2-3), 123-142.

Monsell, S., & Driver, J. (Eds.). (2000). *Control of cognitive processes: Attention and performance XVIII.* Cambridge, MA: The MIT Press

Mottron, L., Dawson, M., Soulières, L., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders, 36*, 27-43.

Murray, D. J. (1967). Role of speech responses in short-term memory. *Canadian Journal of Psychology, 21*(3), 263-263.

Owen, A. M., Downes, J. J., Sahakian, B. J., Polkey, C. E., & Robbins, T. W. (1990). Planning and spatial working memory following frontal-lobe lesions in man. *Neuropsychologia, 28*(10), 1021-1034.

Ozonoff, S., Cook, I., Coon, H., Dawson, G., Joseph, R. M., et al. (2004). Performance on Cambridge Neuropsychological Test Automated Battery subtests sensitive to frontal lobe function in people with autistic disorder: Evidence from the Collaborative Programs of Excellence in Autism network. *Journal of Autism and Developmental Disorders, 34*, 135–150.

Phillips, L. H., Wynn, V., Gilhooly, K. J., Della Sala, S., & Logie, R. H. (1999). The role of memory in the tower of london task. *Memory, 7*(2), 209-231.

Plaisted, K., O'Riordan, M., & Baron-Cohen, S. (1998). Enhanced discrimination of novel, highly similar stimuli by adults with autism during a perceptual learning task. *Journal of Child Psychology and Psychiatry, 39*, 765–775.

Russell, J., Jarrold, C., & Henry, L. (1996). Working memory in children with autism and with moderate learning difficulties. *Journal of Child Psychology and Psychiatry, 37*(6), 673-686.

Russell, J., Jarrold, C., & Hood, B. (1999). Two intact executive capacities in children with autism: Implications for the core executive dysfunctions in the disorder. *Journal of Autism and Developmental Disorders, 29*(2), 103-112.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London B Series, 298*(1089), 199-209.

Schunn, C. D., & Reder, L. M. (1998). Strategy adaptivity and individual differences. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 38, pp. 115–154). San Diego, CA: Academic.

Snodgrass, J. G., & Vanderwart, M. (1980). Standardized set of 260 pictures - norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology-Human Learning and Memory, 6*(2), 174-215.

Sokolov, A. N. (1968/1972). Inner speech and thought. New York: Plenum Press, Original work published 1968.

Stewart, M. E., & Austin, E. J. (2009). The structure of the autism-spectrum quotient (AQ): Evidence from a student sample in Scotland. *Personality and Individual Differences, 47*(3), 224-228.

Tam, H., Jarrold, C., Baddeley, A. D., & Sabatos-DeVito, M. (2010). The development of memory maintenance: Children's use of phonological rehearsal and attentional refreshment in working memory tasks. *Journal of Experimental Child Psychology, 107*(3), 306-324.

Thorndike, E. L., & Lorge, I. (1944). *Tire teacher's word book of 30,000 words.* New York: Teachers' College, Columbia University.

Vygotsky, L. S. (1987). Thinking and speech. *The collected works of Lev Vygotsky* (Vol. I). New York: Plenum (Original work published in 1934).

Wallace, G. L., Silvers, J. A., Martin, A., & Kenworthy, L. E. (2009). Brief report: Further evidence for inner speech deficits in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(12), 1735-1739.

Ward, G., & Allport, A. (1997). Planning and problem-solving using the five-disc Tower of London task. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology, 50*(1), 49-78.

Wechsler, D. (2000) *Wechsler Adult Scale of Intelligence – Third Edition UK*. San Antonio, TX: The Psychological Corporation.

Whitehouse, A. J. O., Maybery, M. T., & Durkin, K. (2006). Inner speech impairments in autism. *Journal of Child Psychology and Psychiatry, 47*(8), 857-865.

Williams, D., Happe, F., & Jarrold, C. (2008). Intact inner speech use in autism spectrum disorder: Evidence from a short-term memory task. *Journal of Child Psychology and Psychiatry, 49*(1), 51-58.

Williams, D. M., & Jarrold, C. (2010). Brief report: Predicting inner speech use amongst children with autism spectrum disorder (ASD): The roles of verbal ability and cognitive profile. *Journal of Autism and Developmental Disorders, 40*(7), 907-913.

Winsler, A., Abar, B., Feder, M. A., Schunn, C. D., & Rubio, D. A. (2007). Private speech and executive functioning among high-functioning children with autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 37*, 1617-1635.

Winsler, A., de Leon, J. R., Wallace, B. A., Carlton, M. P., & Willson-Quayle, A. (2003). Private speech in preschool children: Developmental stability and change, across-task consistency, and relations with classroom behaviour. *Journal of Child Language, 30*, 583-608.

Winsler, A., & Naglieri, J. (2003). Overt and covert verbal problem-solving strategies: Developmental trends in use, awareness, and relations with task performance in children aged 5 to 17. *Child Development, 74*(3), 659-678.

Winsler, A., Fernyhough, C. & Montero, I. 2009. *Private speech, executive functioning, and the development of verbal self-regulation.* New York: Cambridge University Press

Woodbury-Smith, M. R., Robinson, J., Wheelwright, S., & Baron-Cohen, S. (2005). Screening adults for asperger syndrome using the AQ: A preliminary study of its diagnostic validity in clinical practice. *Journal of Autism and Developmental Disorders, 35*(3), 331-335.

World Health Organization. (1992). ICD-10 :The ICD-10 classification of mental and behavioural disorders ; clinical descriptions and diagnostic guidelines. Geneva: World Health Organization.

Table 1: Participant characteristics for Experiment 1

	Group				
	ASD (*n* = 17)	Comparison (*n* = 17)	*t*	*p*	*Cohen's d*
Age	42.13 (14.14)	39.43 (12.51)	0.59	.56	0.20
VIQ	112.82 (11.84)	117.59 (13.13)	-1.11	.28	0.38
PIQ	112.88 (15.33)	112.59 (11.05)	0.06	.95	0.02
FSIQ	114.00 (13.39)	116.71 (13.32)	-0.59	.56	0.20
AQ	33.88 (7.05)	12.13 (5.86)[a]	9.60	<.001	3.36
ADOS Total	10.00 (3.46)	-	-	-	-

[a]Based on 16/17 comparison participants

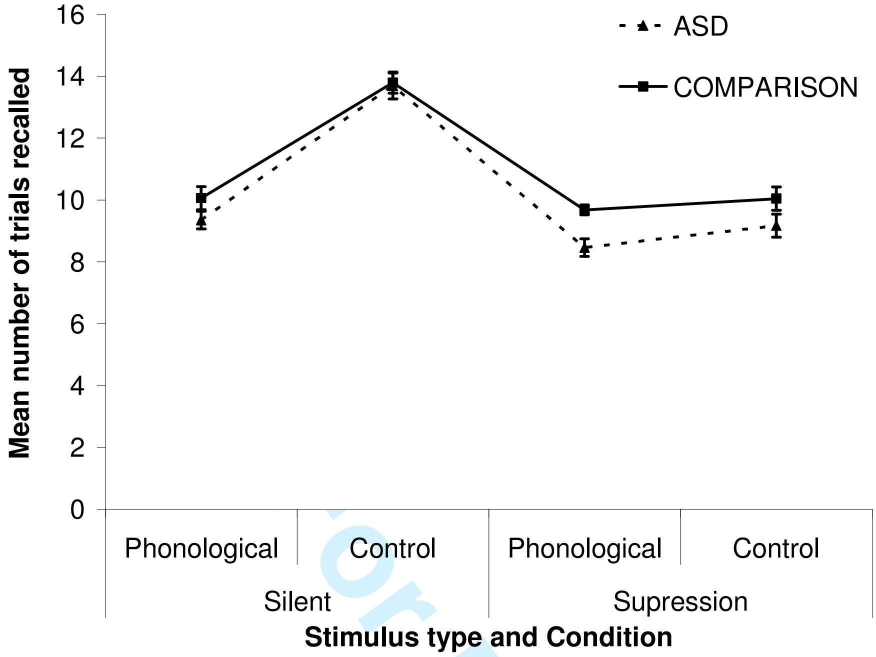
Table 2: Mean (S.D.) number of phonological and control trials recalled by ASD and comparison participants in each condition of Experiment 1

Condition	Stimulus type	Group	
		ASD (*n* = 17)	Comparison (*n* = 17)
Silent	Phonological	9.35 (2.31)	10.06 (3.10)
	Control	13.69 (3.46)	13.79 (2.78)
Articulatory suppression	Phonological	8.46 (2.34)	9.68 (2.54)
	Control	9.17 (3.06)	10.04 (3.13)

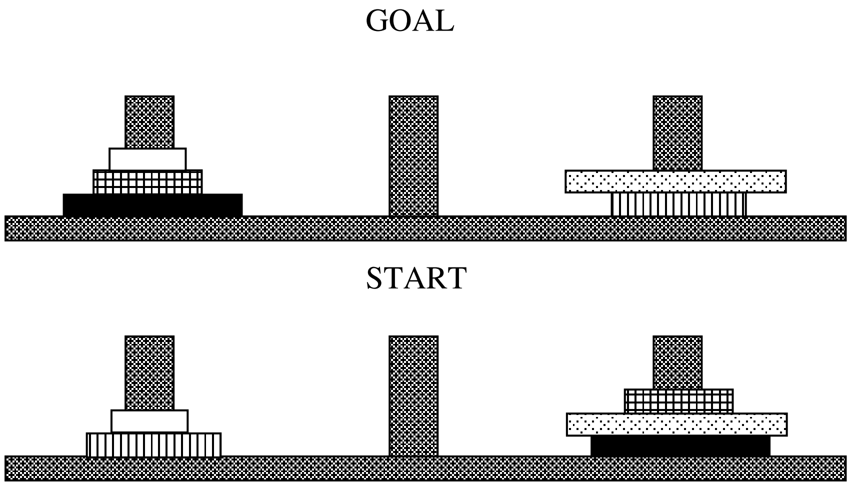
Figure Captions

Figure 1: Serial recall performance on each type of trial (phonological, control) in each condition (silent/suppression), among ASD and comparison participants in Experiment 1. Errors bars represent 1 *SEM*

Figure 2: Example of materials from Experiment 2. The trial displayed takes a minimum of 9 moves to solve (actual disk colours were red, green, yellow, blue, & white).



16
14
12
10
8
6
4
2
0
Mean number of trials recalled
ASD
COMPARISON
Phonological
Control
Phonological
Control
Silent
Supression
Stimulus type and Condition


GOAL
START

Footnotes

1. One participant in the ASD group scored above the defined cut-off for ASD on the "Reciprocal Social Interaction" subscale of the ADOS, but not on the "Total Social Interaction + Communication" scale. We included this participant because they had received a diagnosis of Asperger's syndrome from a leading UK assessment team and because they scored 42 on the AQ, well above the cut-off for ASD on this measure.

2. The exact same pattern of results was observed if other criteria were employed for determining categorically whether an articulatory suppression effect was displayed by participants. For example, if an articulatory suppression effect was defined as taking ≥ 5 moves to complete the Tower puzzles in the articulatory suppression condition than in the silent condition (which is equivalent to a drop in performance across conditions of approximately 1 *SD*), 75% of comparison participants displayed the effect, compared to only 26.7% of participants with ASD, $\chi^2 = 7.24$, $p = .007$, $\varphi = .48$.

3. The subscales of the AQ were determined *a priori* by Baron-Cohen et al. (2001). Three subsequent studies have explored the structure of the AQ using factor analysis. Two studies converge upon a three factor structure ("Social skills", "Communication/mindreading", "Details/patterns"; Austin, 2005; Hurst, 2007) and one study suggests a four factor structure ("Social skills", "Communication/understanding others", "Patterns", "Imagination"; Stewart & Austin, 2009). Notably, all studies contain a "Social skills" subscale, as well as a "Communication" subscale.

Correlation analyses were re-run using the scores from the AQ subscales suggested by each of the factor analytic studies. Results were identical to the original analyses, with communication skills being related uniquely to inner speech use among participants with ASD. The size of the articulatory suppression effect was associated significantly both with the "Communication/understanding others" subscale suggested by Stewart and Austin (2003), $r_s = .74$, $p = .002$, and with the "Communication/mindreading" subscale suggested by Austin (2005) and Hurst (2007), $r_s = .73$, $p = .002$. No other correlations were significant (all other r_ss < .47, all *p*s > .07).

Across all *three* factor analytic studies, six questions from the AQ consistently loaded significantly onto a Communication factor. We ran an additional correlation analysis, exploring the association between the size of the articulatory suppression effect and the communication score derived from only these six questions. Remarkably, this association was also highly significant, $r_s = .73$, $p = .002$. Across all *four* studies of the AQ, including the original study by Baron-Cohen et al. (2001), three questions from the AQ consistently loaded significantly onto a Communication factor. The association between the size of the articulatory suppression effect and the score derived from only these three questions was highly significant, $r_s = .85$, $p < .001$