Citation: Williams, D., Boucher, J., Lind, S. E. & Jarrold, C. (2013). Time-based and event-based prospective memory in autism spectrum disorder: The roles of executive function and theory of mind, and time estimation. Journal of Autism and Developmental Disorders, 43(7), pp. 1555-1567. doi: 10.1007/s10803-012-1703-9

Title: Time-based and event-based prospective memory in autism spectrum disorder: The roles of executive function, theory of mind, and time estimation

David Williams, Department of Psychology, Durham University, Durham, DH1 3LE, UK

Jill Boucher, Department of Psychology, City University London, EC1V 0HB

Sophie Lind, Department of Psychology, Durham University, Durham, DH1 3LE, UK

Christopher Jarrold, Department of Experimental Psychology, Bristol University, Bristol, BS8 1TU

Corresponding author contact details: email: david.williams@durham.ac.uk; Tel: +44 (0)191 33 43270; Fax: +44 (0)191 33 43241

Abbreviated title: Prospective memory in ASD

Full reference: Williams, D. M., Boucher, J. M., Lind, S. E., & Jarrold, C. (2013). Time-based and event-based prospective memory in autism spectrum disorder: The roles of executive function, theory of mind, and time estimation. *Journal of Autism and Developmental Disorders, 43(7),* 1555-1567. doi: 10.1007/s10803-012-1703-9

Abstract

Prospective memory (remembering to carry out an action in the future) has been studied relatively little in ASD. We explored time-based (carry out an action at a pre-specified time) and event-based (carry out an action upon the occurrence of a pre-specified event) prospective memory, as well as possible cognitive correlates, among 21 intellectually high-functioning children with ASD, and 21 age- and IQ-matched neurotypical comparison children. We found impaired time-based, but undiminished event-based, prospective memory among children with ASD. In the ASD group, time-based prospective memory performance was associated significantly with diminished theory of mind, but not with diminished cognitive flexibility. There was no evidence that time-estimation ability contributed to time-based prospective memory impairment in ASD.

Keywords: Autism; prospective memory; theory of mind; executive functioning; cognitive flexibility; set-shifting; time-perception

Abbreviations: Prospective memory (PM).

Time-based and event-based prospective memory in autism spectrum disorder: The roles of executive function, theory of mind, and time-estimation

Autism spectrum disorder (ASD) is a developmental disorder diagnosed on the basis of significant behavioural impairments in social interaction, communication, and behavioural flexibility (American Psychiatric Association, 2000; World Health Organization, 1993). A substantial body of research has documented the profile of abilities and impairments in retrospective memory (i.e., memories or forms of learning acquired from past experiences) among individuals with ASD (e.g., Ben Shalom, 2003; Boucher, Mayes, & Bigham, 2012; Lind, 2010). However, very little research into ASD has explored “prospective memory” (PM), which is the ability to remember to carry out a planned action in the future. More specifically, PM is the ability to execute an intended action after a delay, without any explicit instruction to do so (Kliegel, McDaniel, & Einstein, 2008). Everyday examples include remembering to turn off the bath taps before the bath overflows, remembering to pay a bill on time, or remembering to keep an appointment. A critical distinction drawn by researchers in the field is between “event-based” and “time-based” PM (Einstein & McDaniel, 1990; Einstein, Richardson, Guynn, Cunfer, & McDaniel, 1995; Harris & Wilkins, 1982). Event-based PM involves remembering to carry out an intention upon the occurrence of a particular, specified event (e.g., remembering to remove a pan from the stove *when the timer goes off*). In contrast, time-based prospective memory involves remembering to execute an intention at a particular time-point (e.g., remembering to remove a pan from the stove *in 10 minutes time*). Event-based PM therefore involves the *cued* recall of an intention to do something in the future, whereas time-based PM relies on the *self-initiated* reactivation of an intention, when no external cues or prompts are available. For this reason, event-based PM is thought to require fewer executive resources than time-based PM.

Research on PM among older people has revealed a consistent pattern of decline with age. In a review of the literature, Henry, MacLeod, Phillips, and Crawford (2004) concluded that, in laboratory settings, younger adults tend to show superior event- and time-based PM relative to older adults. However, performance differences between younger and older adults are significantly larger on time-based than event-based tasks. Given that time-based tasks entail a high executive load, this pattern of performance may be attributable to the loss of frontal lobe tissue seen in normal aging (Raz, 2000). Conversely, age-related improvements in PM among typically developing (TD) children, especially on time-based tasks (Kerns, 2000; Ward et al., 2005) may be associated with gains in executive control associated with increasing frontal connectivity (e.g., Passler, Isaac, & Hynd, 1985).

Given the likely role of executive functions in PM, and given the fact that executive function is well known to be impaired in ASD (Hill, 2004; Kenworthy et al., 2008), there is reason to predict that time-based PM will be diminished in people with ASD as a result of executive dysfunction. It is less clear, however, whether event-based PM will also be impaired as a result of executive dysfunction. The executive load in time-based tasks is considered high because *self-initiated* retrieval is required, whereas the executive demands of event-based tasks are comparatively low, because of the *cued* nature of intention retrieval. Importantly, tests of *retrospective* memory show that whereas self-initiated retrieval is impaired across the autism spectrum, cued retrieval is remarkably robust (Boucher et al., 2012; Bowler et al., 2004). It is plausible to predict on these grounds that event-based PM will be intact in ASD, given the cued nature of intention retrieval. In the current study, we employed an independent measure of set-shifting, which allowed us to explore the association between PM performance and degree of cognitive flexibility.

There is a second, and quite different, reason why time-based PM might be relatively more impaired than event-based PM among people with ASD. This relates to a potential difficulty that people with ASD have in perceiving the passage of time and, as a result, with estimating temporal duration. Problems with time perception and estimation in ASD are commonly reported anecdotally and in the clinical literature (for a reviews of this evidence, see Allman & DeLeon, 2011, Boucher, 2001), and have recently been observed in ASD experimentally (at least when very short temporal intervals are involved; see Maister & Plaisted-Grant, 2011; Martin, Poirier, & Bowler, 2010; Szelag et al., 2004). In both the real world and in laboratory settings, time-based PM frequently involves checking a clock or a watch to see whether the specified time interval has elapsed. In studies of PM, clock-checking is generally considered to reflect the participants' efforts to monitor the environment for the PM target (i.e., the time at which they should carry out the intended action). Nonetheless, even when clock-checking is utilised for monitoring, the individual must *estimate when* to check the time. Among individuals with intact time perception/estimation, the frequency of clock-checks is likely to be at its highest as the target time approaches, because the individual has a good sense that time has passed, that the target time must be therefore approaching, and that clock-checking is therefore required. However, a failure to check the time at the appropriate point, because of a diminished sense of the passing of time, would result in PM failure even in a person who had otherwise encoded and stored PM instructions, and who could switch mind-set flexibly when necessary. In the current study, we did not employ an independent measure of time perception/estimation. However, we were nonetheless able to assess the idea that time estimation skills contribute to anticipated time-based PM difficulties in ASD by exploring the pattern of clock checking behaviour made by participants from each group. Diminished time perception could lead to significantly *fewer* clock-checks being made by participants with ASD than by comparison participants because participants with ASD experience time as passing less quickly than comparison participants and hence do not feel an urgency to check the clock. Alternatively, participants with ASD might make a relatively *greater* number of clock-checks, because they experience time as passing relatively more quickly and hence feel a more constant urgency to check the time. Either way, if diminished time-perception contributes to impaired time-based PM in ASD, individuals from the ASD group should show a pattern of clock checks that is distributed evenly over the entire period, rather than showing a peak of clock checks as the target time approaches.

A third factor that might have a negative effect on both event-based *and* time-based PM in ASD is mentalising ability (i.e., the ability to attribute mental states to self and others in order to explain and predict behaviour), because PM of all kinds appears to require the retrieval of a prior *intention*. Given that intentions are an epistemic mental state/propositional attitude and that individuals with ASD have undoubted theory of mind (ToM) impairments (e.g., Happé, 1995), PM of both kinds could be problematic for people with this disorder. Particularly relevant is the finding that individuals with ASD have as much difficulty recognising their own mental states, including their own intentions, as they do recognising mental states in other people (Williams & Happé, 2009, 2010; see Williams, 2010). Therefore, independent of any executive dysfunction in ASD, this deficit in self-awareness of mental states might plausibly result in PM difficulties. The suggestion that PM may be impaired in ASD as a result of reduced ability to represent one's own mental states receives support from studies of neurotypical adults (den Ouden, Frith, Frith, & Blakemore, 2005), and of neurotypical children (Ford, Driscoll, Shum, & Macaulay, 2012). One recent suggestion is that PM is facilitated by engagement in what might be termed "future mentalizing". The idea here is that, at the point of encoding one's intention to perform a particular act in the future, mentally projecting oneself forward in time to imagine the mental state one will be in *at the time* of intention retrieval will increase the likelihood that actual later retrieval will be successful (Ford et al., 2012). This is an interesting idea, linking PM, theory of mind, and "episodic future thinking", each of which may be impaired in ASD (see Lind & Bowler, 2011). However, there may be a more direct link between PM and theory of mind that does not implicate future thinking. Specifically, when the point comes to carry out one's planned action, people with ASD may have difficulty in retrieving and recognising their previously-formed intention simply because it is a mental state. In the example of taking a pan off a stove, whereas neurotypical individuals may relatively easily grasp that they had *meant* to take the pan off the stove, individuals with ASD may simply not recognise such an intention even though such an intention had been formed. As highlighted above, recognition of own present intentions, let alone past or future intentions, is impaired in ASD (Williams & Happé, 2010; Williams, 2010). This alone could result in a significant deficit in PM. In the current study, we employed an independent measure of mentalising ability to explore the association between PM performance and theory of mind.

Before we go on to discuss previous studies of PM in ASD, three methodological points need to be made. First, it is essential that participant groups are *closely*[1] matched for verbal and non-verbal ability, as well as for chronological age. Any differences in experimental task performance between inadequately matched groups could merely be the result of differences in baseline abilities and not true differences in the specific ability/trait being assessed by the experimental task. Second, retrospective memory for PM task instructions must be assessed after completion of the PM task itself. PM failure by an individual who had not encoded and stored the PM task instructions in retrospective memory would, clearly, have a different underlying cause than PM failure by an individual who had encoded and stored the PM instructions, but simply failed to act at the appropriate point. Only in the latter example should we conclude that PM failure was due to diminished functioning of the cognitive mechanism(s) that underlie PM uniquely. For this reason, the majority of studies exploring PM among neurotypical individuals exclude participants who cannot recall the PM instructions. Third, performance on the ongoing task in which the PM targets are embedded needs to be considered when interpreting experimental PM task performance. If participant groups are not equated for ongoing task performance, then group differences in performance of the PM task might result from differences in the ratio of attentional resources devoted to the ongoing task relative to the PM task[2].

To our knowledge, only one study has explored *both* event-based and time-based PM among the same group of individuals with ASD (Altgassen, Koban, & Kleigel, 2012). This study assessed a relatively large sample of intellectually high-functioning *adults* with ASD ($n = 25$) and neurotypical comparison adults ($n = 25$) using a naturalistic ongoing task (preparing breakfast) in which two event-based and two time-based PM tasks were embedded. In this naturalistic task, participants with ASD carried out event-based and time-based PM instructions significantly less reliably than comparison participants. Moreover, participants with ASD also performed significantly less well than comparison participants on an additional laboratory test of event-based PM. Taken together, these findings suggest an across-the-board difficulty with PM in ASD. However, there are reasons for caution in interpreting these findings. In particular, participants with ASD performed markedly less well than comparison participants on the ongoing naturalistic task, completing significantly fewer of the pre-defined ongoing tasks, adhering less frequently to pre-defined rules, monitoring time less frequently, and showing less efficiency overall in carrying out their plans. This suggests a pervasive difficulty with completing the task as a whole, and raises the question of whether participants with ASD fully encoded and stored the task instructions, especially given that participants' retrospective memory for instructions was not checked after the task had been completed.

Three other studies have directly explored PM in ASD using traditional laboratory-based tasks. Two of these studies assessed event-based PM, only: one reporting a significant diminution of this ability among participants with ASD (Brandimonte, Filippello, Coluccia, Altgassen, & Kliegel, 2011) and one reporting undiminished performance in ASD (Altgassen, Schmitz-Hübsch, & Kliegel, 2010). A third study assessed time-based PM, only, reporting significant impairment in participants with ASD (Altgassen, Williams, Bölte, & Kliegel, 2009). There are potential methodological concerns about these studies, however. Specifically, in neither the study by Brandimonte et al. (2011) nor the study by Altgassen et al. (2009) were participant groups whose PM performance was analysed well-matched for verbal mental age (Altgassen et al., 2009) or IQ (Brandimonte et al.). Moreover, in the Altgassen et al. (2009) study, the ASD group performed considerably less well than comparison participants on the ongoing task, again calling into question the reliability of any finding of impaired PM *per se.*

Moreover, only one of the studies of PM in ASD, that by Altgassen et al. (2012), has attempted to assess the cognitive correlates of PM in ASD. However, only correlations among collapsed diagnostic groups were reported. This is not entirely appropriate, because there is no reason to believe that the cognitive correlates of PM among individuals with ASD will be necessarily the same as the cognitive correlates of PM among neurotypical individuals. As argued above, a range of ASD-specific factors might be likely to impair PM in ASD, notably problems of cognitive flexibility, difficulty in estimating temporal duration, and impaired mentalising ability.

In sum, there is a need for further studies of PM in ASD, preferably testing the same participants on both event-based and time-based tasks, and observing the key methodological points noted above. The first aim of the study reported here is to respond to this need. We predict that time-based PM will be impaired in ASD. Whether or not event-based PM will also be impaired is an open question, given the difficulties with interpreting the results of previous studies of this ability in ASD.

The second aim was to investigate the cognitive correlates of PM in ASD, to facilitate understanding of the underlying basis of any impairment(s) that might be demonstrated. If we were to demonstrate impairment in one or both PM conditions, then understanding the underlying basis of such impairment(s) would contribute to the development of strategies for facilitating the ability of people with ASD to ‘remember to remember’ in everyday life. Among participants with ASD, significant associations between time-based PM performance, and both cognitive flexibility and theory of mind were predicted. A significant association between event-based PM and theory of mind was also expected. However, because of the relatively low executive load inherent to event-based PM tasks, we did not necessarily expect to see an association between event-based PM task performance and cognitive flexibility. In order to investigate the influence of time estimation skills on time-based PM ability, we explored the patterns of clock-checking behaviour displayed by participants from each group, predicting a group difference in the frequency of clock-checks and a qualitatively different pattern of checking among participants with ASD.

Method

Participants

Ethical approval for this study was obtained from the appropriate University Ethics Committee. Twenty one children with ASD and 21 TD comparison children took part in this experiment, after their parents had given written, informed consent. Participants in the ASD group had all received formal diagnoses of autistic disorder (n = 13) or Asperger's disorder (n = 8), according to conventional criteria (American Psychiatric Association, 2000; World Health Organisation, 1992). To assess ASD features, parents of participants with ASD completed the Social Responsiveness Scale (SRS; Constantino, Davis, Todd, Schindler, Gross, Brophy et al., 2003). In addition, among those parents of ASD participants who agreed to be contacted over the phone, 10 also completed the Developmental, Dimensional, and Diagnostic interview (3Di; Skuse et al., 2004). In each case, participants' scores were in the range expected for individuals with a diagnosis of ASD. Parents of comparison children also completed the SRS. All participants in the comparison group scored below the defined cut-off for ASD on the SRS.

Using the Wechsler Abbreviated Scale of Intelligence (WASI; Psychological Corporation, 1999), the groups were equated closely for verbal and non-verbal ability. The groups were also equated closely for chronological age. Participant characteristics are presented in Table 1.

Table 1 here

Tests and procedures

Participants were tested individually either in a sound-attenuated laboratory at the university at which the research was based, or in a quiet room in the participant's school. Testing was conducted over three sessions, usually completed two to three days apart, each lasting approximately 45 minutes. The IQ test was always administered first. Frequently, one of the PM conditions was completed next, followed by the test of cognitive flexibility, followed by the ToM task, followed by the remaining PM condition. However, there was no fixed, invariant order to task presentation (although presentation of conditions *within* tasks was always fully counterbalanced across participants).

Prospective memory tasks

The ongoing task for both the event-based and time-based PM measures was a 2D computer-based driving game, presented to participants on a ThinkPad laptop. The task was based on that employed by Kerns (2000) (see Figure 1). The aim of the task was to drive a car down a street while collecting gold tokens and avoiding obstacles in the form of other cars, barriers and crates. Players were awarded five points per token that they collected, and deducted five points for every hazard that they hit. Participants controlled the game using a Saitek games controller, allowing the player to turn left/right, and move forward and backwards. Participants had to press the forward arrow on the controller to move the car and increase speed. If the car was moving at speed and no keys were pressed the car slowed down gradually until it stopped. The ongoing task lasted for 8 minutes in both conditions.

Verbal instructions for both the ongoing task and for the PM component of the task in each condition (detailed below) were given by the experimenter prior to testing. The experimenter then gave a one and a half minute demonstration of the ongoing task including successful performance of the relevant PM task. In addition, immediately before testing in each condition, an on-screen cartoon mechanic, “Mike the Mechanic”, appeared and reiterated (via text written in a speech bubble) the relevant PM instructions (as detailed next), as well as the ongoing task demands. The order in which participants completed the time-based and event-based conditions of the task was counterbalanced across participants.

In the time-based condition, participants were told that the car had a limited amount of fuel, and that this would run out in 80 seconds unless replenished. The PM task was to replenish the fuel before it ran out. The fuel level could be checked by pressing the ‘F’ key on the laptop, in response to which a fuel gauge appeared in the bottom left-hand side of the screen. The gauge appeared for three seconds, after which it disappeared. When the fuel level was high (during the first 60 seconds of driving) the gauge level was green and participants could not refuel. When the fuel level was low (between 60 and 80 seconds) the gauge level changed from green to red. While the fuel gauge was red participants could refuel the car by pressing the ‘R’ key. Hence, there was a 20 second window for participants to refuel the car. If participants ran out of fuel the car stopped, their ongoing task score was reset to zero, and the fuel gauge was filled up to maximum. A reminder from Mike the mechanic also appeared on the screen that read “Remember not to run out of fuel!”. Participants had the opportunity to refill the car on a maximum of six occasions. Running out of fuel was considered a PM failure. Measures recorded were: number of PM failures; on-going task score (number of points gained from collecting gold tokens minus number of points deducted for colliding with hazards); total number of fuel checks; and the temporal distribution of fuel checks. The distribution of fuel checks was measured by noting the number of checks made by participants across each minute of the trial, broken down into 15 second time periods between refuellings (i.e., 0-15 seconds after refuelling, 16-30 seconds after refuelling, 31-45 seconds after refuelling, and 46-60 seconds after refuelling). This breaking down of trials into intervals to measure the distribution of fuel checks is typical in studies of time-based PM (e.g., Kerns; 2000; Mackinlay, Kliegal, & Mäntyla, 2009; Mäntyla, Carelli, & Forman, 2007).

In the event-based condition, the PM instruction was to press the 'H' key every time the car they were driving in the ongoing task passed a lorry. A lorry, which participants needed to manoeuvre around, appeared every 80 seconds. During this version of the task the car had an unlimited amount of fuel and nothing happened if participants pressed either the 'F' key or the 'R' key. Six lorries appeared during the 8-minute ongoing task. Failing to press the required key within 5 seconds of passing a lorry was considered a PM failure. Measures recorded were: number of PM failures; and ongoing task score (number of points gained from collecting gold tokens minus number of points deducted for colliding with hazards). With regard to the ongoing task score, it is important to note that the measure was independent of PM task performance. Although a failure in the time-based PM condition (i.e., running out of fuel) resulted in the ongoing task score being reset to zero, the score we used to analyse ongoing task performance did not include any resets; it was merely the number of points gained from collecting gold tokens minus number of points deducted for colliding with hazards, which was recorded directly by the computer.

At the end of each PM task, retrospective memory for PM and ongoing task instructions was assessed. Participants were asked "What did you have to do in the game? What did Mike the mechanic tell you to do at the start of the game?" Participants were required to recall the specific PM task instruction, including the specific keyboard keys it was necessary to press, as well as the fact they needed to collect coins and avoid hazards. Each and every participant in the current study recalled these instructions, indicating that they had successfully encoded and stored, and were able to retrieve, task details.

Figure 1 here

Tests of possible cognitive correlates

As noted, three possible cognitive correlates of PM performance were explored in this study: cognitive flexibility, time estimation, and theory of mind. Time estimation was assessed using information that emerged directly from within the time-based PM task, namely the total number of fuel checks made during the time-based PM task and the distribution of checks across the duration of the task (see above). The argument behind the use of these measures was that impaired temporal estimation would be associated with an abnormal total number of checks, and with an atypical tendency to check consistently throughout the task, rather than showing increased checking only as the target time approached. Cognitive flexibility and theory of mind were assessed using additional measures, specifically the Wisconsin Card Sorting Task (WCST; Berg, 1948) and the Animations Task (Abell, Happé, & Frith, 2000) respectively.

The *WCST* is a classic measure of “set-shifting”, or “cognitive flexibility” (Miyake et al., 2000). In the current study, we employed a modified version of the WCST that was introduced by Heaton (1976). The modified WCST is designed to reduce administration time and difficulty level. It is sensitive to frontal lobe pathology among neuropsychological patients (e.g., Nelson, 1976), and is used widely as a measure of cognitive flexibility among neurotypical children (e.g., Cianchetti et al., 2007), and neurotypical adults (e.g., Obonsawin et al., 2002). The modified WCST comprises four stimulus cards and 48 response cards that vary on three dimensions: type of shape (triangle, cross, circle, star), number of shapes (one to four), and colour of shapes (red, green, blue, yellow). Response cards that share more than one attribute (e.g., shared number and colour) are not included. The four stimulus cards are placed separately, face-up in front of the participant, with the response cards in a pack, face down. The participant’s task is to turn over the response cards one by one, placing each card below the stimulus card it ‘matches’ – i.e. the task is to sort the response cards into categories according to one of these dimensions, as displayed in the stimulus cards. For each participant response, the experimenter provides positive or negative feedback, but does not tell the participant explicitly what the (arbitrarily) ‘correct’ sorting strategy is. Thus, participants must infer the sorting strategy from the experimenter’s feedback. After 6 consecutive cards have been sorted correctly, the experimenter tells the participants that the rule has now changed, and that they must sort the remaining response cards utilising a different rule.

Key measures of performance on the WCST are number of perseverative errors and number of non-perseverative errors. An error was scored as perseverative when participants a) persisted in sorting a card into the same category as the previous sort, after being told their previous card sort was incorrect; b) continued to sort cards using the previously correct rule, after they had been told that the rule had changed; and c) continued to sort according to the previously completed category (see Cianchetti et al., 2007).

The *Animations* task requires participants to describe interactions between a large red triangle and a small blue triangle, as portrayed in a series of silent video clips. Four clips (taken directly from Abell et al., 2000) were employed. On the one hand, two of the clips were designed such that explanation of the triangles' behaviour required the attribution of epistemic mental states, such as belief, intention, and deception ("Coaxing" and "Tricking" animations). On the other hand, the remaining two clips were designed such that explanation of the triangles' behaviour required the attribution of physical agency, without reference to complex mental states ("Fighting" and "Following" animations). Following Heider and Simmel (1944), Abell et al. referred to these clips as involving "goal-directed" action (e.g., copying, chasing, following). While these are certainly goal states, and hence could be considered mental states (albeit non-epistemic mental states), the explanation of the triangles' behaviour in these clips requires only a focus on the *actions* displayed by the characters, rather than on the underlying *mental states* that cause the actions. For example, to describe two characters as "kissing" does not require any understanding of the mental states underlying that behaviour, although it does require the perception of the characters as animate agents. It is for this reason that we refer to these clips as comprising a "Physical" condition, which we contrast with the Mentalising condition, described above.

Each clip was presented to participants on a computer screen. To familiarise participants with the task, two practice animations were shown before the experimental stimuli. Participants were asked to describe the behaviour displayed in each of these clips, and experimenter feedback was given after each description. For the experimental animations, participants watched each clip twice. First, participants watched the clip through once in silence and were told to "watch the clip and see how the triangles are interacting". Participants then watched the clip again and were told "as you watch the clip again I would like you to tell me how the triangles are interacting". Participants provided a running commentary on the triangles' interactions during this second presentation of the clip. For the experimental trials, a digital audio recording of participants' responses was made for later transcription. No feedback was given on the experimental trials. The order in which the experimental clips were presented was counterbalanced across participants.

Participants' descriptions for each animation were scored on the basis of scoring criteria outlined in Abell et al. (2000; see their Appendix A for detailed scoring criteria). Participants' descriptions of each animation were given a score of 0, 1, or 2 according to their level of accuracy. Accuracy was defined as the extent to which the participant's description captured the intended meaning of the animation. Thus, the score achievable in each condition (ToM/Physical) was between zero and four. Each description was also scored by an independent rater who was blind to group status. Inter-rater reliability was near-perfect according to Landis and Koch's (1977) criteria**,** $\kappa = .90$.

Results

Prospective memory tasks

Performance on the PM component in both conditions

Table 2 shows the mean number of PM failures made by ASD and TD participants in each condition (time-based/event-based) of the task. A mixed ANOVA was conducted on these data, with Group (ASD/comparison) as the between-participants variable and PM Condition (time-based/event-based) as the within-participants variable. There was a significant main effect of Condition, $F(1, 40) = 21.91$, $p < .001$, reflecting superior performance in the event-based condition than in the time-based condition. The main effect of Group was not significant, $F(1, 40) = 1.57$, $p = .22$. However, there was a significant interaction between Condition and Group, $F(1, 40) = 6.46$, $p = .02$. Independent-samples *t*-tests revealed that children with ASD performed significantly *less well* than comparison participants in the time-based condition, $t(32.65) = 2.20$, $p = .03$, $d = 0.68$, but non-significantly *better* than comparison participants in the event-based condition, $t(40) = 0.56$, $p = .58$, $d = 0.17$. Paired-samples *t*-tests revealed that, among comparison participants, performance in the time-based condition was not significantly different from performance in the event-based condition, $t(20) = 1.45$, $p = .16$, $d = 0.44$. However, among ASD participants, performance in the time-based condition was significantly poorer than performance in the event-based condition, $t(20) = 5.35$, $p < .001$, $d = 1.18$.

Among comparison participants, performance in both the event-based and time-based conditions was significantly above floor and significantly below ceiling (all *p*s < .05). Among ASD participants, performance in both conditions was significantly above floor (all *p*s < .05) and performance in the time-based condition was significantly below ceiling, $t(20) = 5.28$, $p < .001$. Performance in the event-based condition was below ceiling, but marginally non-significantly so, $t(20) = 1.83$, $p = .08$.

Performance on the ongoing component

Table 2 shows overall ongoing task performance collapsed across event-based and time-based conditions among ASD and comparison participants. A mixed ANOVA was conducted on these data, with Group (ASD/comparison) as the between-participants variable and Condition (time-based/event-based) as the within-participants variable. There was a significant main effect of Condition, $F(1, 40) = 10.78$, $p = .002$. This reflected the fact that ongoing task performance was poorer in the time-based condition (M = 321.67, SD = 113.33) than in the event-based condition (M = 357.62, SD = 122.91). Crucially, however, there was no significant main effect of Group, $F(1, 40) = 2.61$, $p = .11$, and no significant interaction between Group and Condition, $F(1, 40) = 0.01$, $p = .93$.

These results are clearly inconsistent with the view that impaired PM among participants with ASD in this study is merely a consequence of a group difference in ongoing task performance. However, to be absolutely stringent, we re-ran the analysis concerning PM failures (see the section immediately above), but this time *covaried* ongoing task performance score. The data met the assumptions for ANCOVA. Essentially, the results from the ANCOVA were the same as from the ANOVA. Further confirming that group differences in PM performance were independent of group differences in ongoing task performance, the critical interaction between Condition and Group remained significant in the ANCOVA, $F(1, 39) = 6.08$, $p = .02$[3].

Table 2 here

Measures of possible cognitive correlates

Fuel checking behaviour

Figure 2 shows the number and distribution of clock-checks made in each of the four 15 second intervals preceding the target time (averaged across trials) of the time-based PM task. A mixed ANOVA was conducted on this these data, with Group (ASD/comparison) as the between-participants variable and Time period (0-15 seconds/16-30 seconds/31-45 seconds/46-60 seconds) as the within-participants variable. There was a significant main effect of Time period, $F(3, 120) = 52.22$, $p < .001$. Within-participant contrasts suggested significant linear ($p < .001$) and quadratic ($p = .005$) effects. The main effect of Group was not significant, $F(1, 40) = 0.09$, $p = .77$. The interaction between Group and Time period was not significant, $F(3, 120) = 1.35$, $p = .26$. The total number of fuel checks made prior to each target time (i.e., the period when the fuel could be refilled) was negatively correlated with the number of PM failures among ASD ($r = .80$, $p < .001$) and comparison participants ($r = .47$, $p = .03$). Thus, there no evidence of a difference between the groups in time estimation, or that a difficulty with time estimation among participants with ASD could have contributed to diminished time-based PM performance.

Figure 2

Wisconsin Card Sorting Test and Animations task

Table 3 shows the performance of ASD and TD participants on the modified WCST and on the Animations test. Participants with ASD showed a largely expected pattern of performance, reflecting diminished cognitive flexibility and diminished mentalising ability. Specifically: relative to comparison participants, participants with ASD made significantly more perseverative (but not non-perseverative) errors on the WCST, and performed significantly less well in the Mentalising condition of the Animations task (but not the Physical condition).

Table 3

Correlational analyses

Given that there were no significant group differences in event-based PM and that performance in this condition was not significantly below ceiling among ASD participants, we restricted our correlation analyses to data from the time-based condition of the PM task. Within each group, we explored the extent to which number of time-based PM failures was associated with performance on the WCST and with performance on the Animations task. We did this by conducting a series of *partial* correlation tests (see Table 4). First, we assessed the association between number of PM failures and number of perseverative errors on the WCST, controlling for number of *non*-perseverative errors on the WCST. Arguably, this provided a pure test of the relation between PM ability and the kind of perserveration that is characteristic of executive dysfunction, independent of general (non-executive) demands inherent to the WCST task. Next, we assessed the association between the number of PM failures and score in the Mentalising condition of the Animations task, controlling for score in the Physical condition of the Animations task. Arguably, this provided a pure test of the relation between PM ability and mentalising ability, independent of general (non-mentalising) demands inherent to the Animations task.

From Table 4 it can be seen that among participants with ASD there was only a relatively modest correlation between number of PM failures and number of perserverative errors on the WCST, which did not approach statistical significance, *p* = .17. On the other hand, there was a moderate-to-large association between number of PM failures and score in the Mentalising condition of the Animations task, *p* = .03. Among comparison participants, a somewhat different picture emerged. PM failure was moderately associated with perseverative errors on the WCST, although the association fell short of statistical significance, *p* = .09. The association between PM failure and Mentalising score on the Animations task ability was small and did not approach statistical significance, *p* = .30.

Next, we addressed whether these correlations with PM accuracy were specific to PM, or whether they merely reflected association with general (non-PM) demands of the task. Thus, we explored whether set-shifting and/or mentalising ability was associated with overall ongoing task performance. Again, we conducted a series of *partial* correlations (see Table 4) using the same approach as described above in relation to number of PM failures, but this time involving ongoing task score. No association reached statistical significance even on one-tailed significance tests (all *r*s ≤ -.32, all *p*s ≥ .09).

We explored further the only statistically significant correlation to emerge from these analyses, which was between number of PM failures and score in the Mentalising condition of the Animations task among ASD participants. First, to be absolutely stringent, we re-ran the analysis controlling for ongoing task score *in addition to* controlling for score in the Physical condition of the Animations task. The association between number of PM failures and score in the Mentalising condition of the Animations task remained significant among participants with ASD, *r* = -.42, *p* = .02 (one tailed). Second, Fisher's Z tests indicated that the association between PM failures and score in the Mentalising condition of the Animations task was significantly larger among ASD than comparison participants, *z* = -1.76, *p* = .04 (one tailed).

Table 4 here

Discussion

Until now, our understanding of the profile of PM abilities/impairments among individuals with ASD has been limited by the fact that the only experimental studies of PM in ASD have arguably had methodological difficulties that have made interpretation of findings difficult. In addition, none of the published studies investigated ASD-specific cognitive correlates of any impairment(s) of PM.

The study reported here is the first to assess both event-based and time-based PM in the same participants, using carefully-controlled laboratory-based PM tasks. Groups of participants in the current study were matched closely for age, as well as for verbal and non-verbal abilities, and all stored and encoded the PM task instructions successfully. Moreover, there were no significant differences between the groups in ongoing task performance. Thus, the current study arguably provides the clearest indication to date of the profile of PM abilities among individuals with ASD. The results were quite clear: In the overall ANOVA on experimental task error data, the interaction between Group (ASD/comparison) and PM task (event-based/time-based) was highly significant, reflecting diminished time-based PM among participants with ASD, but undiminished event-based PM.

With regard to the finding of unimpaired event-based PM, our study is in agreement with Altgassen et al. (2010), but conflicts with Brandimonte et al. (2011) and Altgassen et al. (2012). As highlighted above, there are arguably concerns about each of these latter studies. At best, we can say that there is a mixed pattern of findings from studies that have investigated PM directly. However, a study by Jones, Happé, Pickles, Marsden, Tregay, et al. (2011) may help us to decide which of these discrepant findings is reliable. This study was not discussed above because it did not set out to study PM directly. Rather, Jones et al. set out to explore “everyday memory” among a very large sample of adolescents with ASD (n = 94) and closely-matched comparison participants (n = 55). Participants completed selected tasks from the Rivermead Behavioural Memory Test (RBMT; Wilson & Baddeley, 1985). This naturalistic test assesses various domains of memory functioning across a number of different sub-tests, three of which assess event-based PM. Adopting an approach that is common within the RBMT literature, Jones et al. calculated a composite PM score based on performance across these three subtests and reported a significant event-based PM impairment in ASD (although the effect size was only small; d = 0.40). However, the scoring procedure employed by Jones et al. (2011) conflates retrospective memory and PM by including individuals who could not recall the PM instructions even after prompting[4]. If participants who forgot the PM instructions entirely are excluded from Jones et al.'s results (as would be the case in most traditional studies of PM), there was in fact no hint of any group differences in task performance[5]. Thus, in keeping with our results, as well as those of Altgassen et al. (2010), Jones et al. found that PM task performance was *not* impaired in a large sample of individuals with ASD on a naturalistic event-based task. Given that the current study is the third to observe undiminished event-based PM performance among individuals with ASD, and that there are concerns about the two studies of event-based PM that have observed an ASD-specific impairment, we suggest that this aspect of PM may not be significantly impaired in ASD. However, two other possibilities are evident. First, it could be that our event-based PM task was simply not sensitive enough to detect an ASD-specific impairment in this ability. Contrary to this possibility, however, is the fact that participants with ASD performed non-significantly *better* than comparison participants on the event-based PM task. While comparison participants performed significantly below ceiling on this task, indicating that the task was sensitive enough to detect individual differences in PM ability among neurotypical children, participants with ASD performed at near ceiling levels on the task, suggesting that event-based PM is a strength for people with this disorder. Second, it could be that individuals with ASD were using alternative strategies to *perform* well on the event-based PM task despite limited underlying *competence*. With regard to PM performance among neurotypical individuals, most adults and children report not having consciously thought about the PM instructions until the relevant event occurs (e.g., Ward et al., 2005). This suggests a form of "spontaneous retrieval" of one's intention upon the encountering of the relevant event. It may be that individuals with ASD do not rely on (or are not able to employ) such a spontaneous retrieval strategy to succeed on PM tasks. One possibility is that instead individuals with ASD tend to constantly rehearse the

PM task instructions throughout the task. Certainly, individuals with ASD tend rely on the verbal rehearsal of information as a means of supporting task performance (see Williams, Bowler, & Jarrold, 2012). Note that such a consistent monitoring/rehearsal strategy would be more applicable to event-based PM tasks than time-based PM tasks, because only in the former is there a cue (i.e., the event itself) to tell you *when* to act. Without such a cue, one may simply miss the *appropriate point* to act even though one is constantly rehearsing the PM instructions. Nonetheless, in the absence of direct evidence of this and given that the current findings fit well with the findings of spared cued retrieval in retrospective memory among people with ASD, the tentative conclusion from this study should be that event-based PM is not impaired in this disorder.

In the time-based condition of our study, participants with ASD forgot to refuel the car (i.e., had more PM failures) significantly more often than comparison participants. The difference between the groups on this key experimental variable was associated with an effect size (Cohen's *d*) of 0.68, suggesting that time-based PM is moderately impaired in ASD. If anything, this may be an underestimate of the severity of time-based PM impairment in ASD, because each time participants had a failure (i.e., every time the car ran out of fuel), the ongoing task provided a reminder of the PM task instruction (i.e., not to run out of fuel). Had the task not provided this reminder, it may be that the time-based PM impairment would have been even more pronounced among participants with ASD. Although participants with ASD ran out of fuel (i.e., had more PM failures) significantly more often than comparison participants, they checked the clock as frequently as did comparison participants, and showed similar patterns of time-monitoring behaviour across the course of the task. This suggests that they retained the task instructions in mind, understood that they needed to monitor the time, and had an accurate sense of the passing of time, but simply failed to carry out their plan at the appropriate time.

The study reported here is also the first to investigate possible *ASD-specific* cognitive correlates of any impairment(s) of PM. Our findings on clock-checking behaviour are inconsistent with the suggestion that impaired time-estimation might contribute to an impairment of time-based PM in ASD. According to this account, impaired time-estimation would have manifested in a more even distribution of checks across the total time period in the ASD group, and to a difference in the frequency of checking. Neither of these predictions was supported. This finding, although negative in one respect, provides an interesting addition to the growing literature on time perception in ASD. To date there have been no studies, to the best of our knowledge, of the ability to estimate durations of longer than 45 seconds, almost all of the studies reporting impairments or anomalies using paradigms involving extremely short durations. Our observations in this study might suggest that any time-estimation impairment is restricted to such durations, at least among intellectually high-functioning individuals.

The hypothesis that impaired cognitive flexibility might contribute to impaired time-based PM was also unsupported. Among participants with ASD, no significant correlation was observed between time-based PM and the ability to flexibly shift mindset (i.e., perseveration on the WCST), despite the fact that both abilities were diminished among participants from this group. It may be that the PM instructions were so inherently related to ongoing task activity that relatively little task-switching was required. If this is the case, then participants with ASD should show even greater time-based PM impairment on a task that placed higher demands on set-shifting than the task employed in the current study. One other possibility for the failure to find a significant association between cognitive flexibility and PM is that the WCST is not a sufficiently pure and sensitive measure of cognitive flexibility. For example, some ASD researchers have suggested that successful performance on the WCST requires substantial inhibitory control, independent of set-shifting ability (e.g., Ozonoff & Strayer, 1997). However, factor analytic studies suggest that perseveration on the WCST is uniquely associated with set-shifting ability (e.g., Miyake et al., 2000). Most important, though, we specifically employed a modified version of the WCST in order to minimise the involvement of demands other than cognitive flexibility. As such, our study could be viewed as providing a more accurate reflection of the specific relation between cognitive flexibility and PM than previous studies that have explored the relation using the standard version of the WCST (e.g., Marsh & Hicks, 1998; Martin, Kliegel, & McDaniel, 2003).

The final hypothesis that ToM ability would correlate with PM in both the event-based and time-based condition was partially supported. Specifically, among participants with ASD, number of time-based

PM failures was associated significantly with performance in the Mentalising condition of the Animations task, independent of any association with performance in the Physical condition of the Animations task. This is a particularly novel finding and suggests that further investigation of the possible links between PM and mentalising would be beneficial (cf. Ford et al., 2012). However, one prediction that could follow naturally from this finding is that event-based PM should be equally as impaired as time-based PM among people with ASD, given that both event-based and time-based PM would appear to involve representation of one's own intentions to the same extent. In fact, however, participants with ASD made non-significantly *fewer* event-based PM errors than comparison participants in the current study (cf. Altgassen et al., 2010). Thus, an explanation of the link between ToM and time-based PM, in the face of undiminished event-based PM, is required. One possibility is that the cued nature of event-based PM removes the need to represent one's intention as an intention in order to act appropriately; mental states, such as beliefs and intentions, are described as propositional attitudes because they relate an individual to a proposition that is about some state of affairs. For example, "believe" could be the attitude that I take toward the proposition "the milk is in the fridge". To succeed on a time-based PM task, it may necessary to recognise and retrieve the attitude (i.e., intention) that one takes toward a proposition (i.e., "refuel the car when the fuel tank is running low"). In contrast, the appearance of the lorry in the event-based PM task may cue recall of the proposition directly ("press the key when a lorry appears"), which is sufficient to trigger behaviour automatically, whether or not one additionally represents the attitude one takes toward the proposition. Of course, it will be important to replicate independently this finding of an association between ToM and time-based PM. However, the fact that among neurotypical individuals associations between ToM and PM have been observed at the behavioural (Ford et al., 2012) and neurobiological (den Ouden et al., 2005) levels of description suggest that further investigation of the issue in ASD is warranted.

In sum, the present study provides further evidence that time-based and event-based PM are separable abilities, and that only the former is impaired in ASD. The finding that theory of mind is characteristically diminished in ASD and that the degree of this diminution was associated with the degree of impairment in time-based PM among people with this disorder suggests that these abilities may be linked at the cognitive level. Future studies should examine the potential links among future thinking abilities (which may also be impaired in ASD; Lind & Bowler, 2009; but see Crane, Lind, & Bowler, in press), theory of mind, and PM. Furthermore, future studies should also explore the possibility that individuals with ASD employ compensatory strategies on event-based PM tasks.

References

Abell, F., Happe, F., & Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development, 15*(1), 1-16.

Allman, M. & DeLeon, I. (2011). No time like the present: Time perception in autism. In A. Giordana & V. Lombardi (Eds), *Causes and Risks for Autism* (pp 65-76). Place: Nova Science Publisher.

Altgassen, M., Williams, T. I., Bölte, S., & Kliegel, M. (2009). Time-based prospective memory in children with autism spectrum disorder. *Brain Impairment, 10,* 52-58.

Altgassen, M., Schmitz-Huebsch, M., & Kliegel, M. (2010). Event-based prospective memory performance in autism spectrum disorder. *Journal of Neurodevelopmental Disorders, 2*(1), 2-8.

Altgassen, M., Koban, N., & Kliegel, M. (2012). Do adults with autism spectrum disorder compensate in naturalistic prospective memory tasks? *Journal of Autism and Developmental Disorders, in press*

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4th edition, text revised) (DSM-IV-TR).* Washington DC: American Psychiatric Association

Ben-Shalom, D. (2003). Memory in autism: Review and synthesis. *Cortex, 39,* 1129-1138.

Berg, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *Journal of General Psychology, 39*(1), 15-22.

Boucher, J. (2001). Lost in a sea of time: Time-parsing and autism. In C. Hoerl & T. McCormack (Eds.), *Time and Memory* (pp. 111-135). Oxford: Clarendon Press.

Boucher, J., Mayes, A., & Bigham, S. (2012). Memory in autistic spectrum disorder. *Psychological Bulletin*, *138*, 458-496.

Brandimonte, M. A., Filippello, P., Coluccia, E., Altgassen, M., & Kliegel, M. (2011). To do or not to do? Prospective memory versus response inhibition in autism spectrum disorder and attention-deficit/hyperactivity disorder. *Memory, 19*(1), 56-66.

Cianchetti, C., Corona, S., Foscoliano, M., Contu, D., & Sannio-Fancello, G. (2007). Modified Wisconsin card sorting test: Normative data in children 4-13 years old, according to classical and new types of scoring. *Clinical Neuropsychologist, 21*(3), 456-478.

Constantino, J. N., Davis, S. A., Todd, R. D., Schindler, M. K., Gross, M. M., Brophy, S. L., et al. (2003). Validation of a brief quantitative measure of autistic traits: Comparison of the social responsiveness scale with the autism diagnostic interview-revised. *Journal of Autism and Developmental Disorders, 33*(4), 427-433.

den Ouden, H. E. M., Frith, U., Frith, C., & Blakemore, S. J. (2005). Thinking about intentions. *Neuroimage, 28,* 787-796.

Einstein, G. O. & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology-Learning Memory and Cognition, 16,* 717-726.

Einstein, G. O., Richardson, S. L., Guynn, M. J., Cunfer, A. R., & McDaniel, M. A. (1995). Aging and prospective memory: Examining the influences of self-initiated retrieval-processes. *Journal of Experimental Psychology-Learning Memory and Cognition, 21,* 996-1007.

Ford, R. M., Driscoll, T., Shum, D., & Macaulay, C. E. (2012). Executive and theory-of-mind contributions to event-based prospective memory in children: Exploring the self-projection hypothesis. *Journal of Experimental Child Psychology, 111*(3), 468-489.

Harris, J. E. & Wilkins, A. J. (1982). Remembering to do things: A theoretical framework and an illustrative experiment. *Human Learning, 1,* 123-136.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology, 57*, 243-259.

Henry, J. D., MacLeod, M. S., Phillips, L. H., & Crawford, J. R. (2004). A meta-analytic review of prospective memory and aging. *Psychology and Aging, 19,* 27-39.

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

Kerns, K. A. (2000). The CyberCruiser: An investigation of development of prospective memory in children. *Journal of the International Neuropsychological Society, 6,* 62-70.

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.

Landis, J. R., & Koch, G. G. (1977). Measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Lind, S. E. (2010). Memory and the self in autism: A review and theoretical framework. *Autism 14*, 430-456.

Mackinlay, R., Charman, T. & Karmiloff-Smith, A. (2006). High-functioning children with autism spectrum disorder: a novel test of multitasking. *Brain & Cognition, 61*, 14-24.

Maister, L. & Plaisted-Grant, K. (2011). Time perception and its relationship to memory in Autism Spectrum Conditions. *Developmental Science, 14,* 1311-1322. DOI: 10.1111/j.1467-7687.2011.01077.x

Marsh, R.L., & Hicks, J.L. (1998). Event-based prospective memory and executive control of working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 336-349.

Martin, M., Kliegel, M., & McDaniel, M.A. (2003). The involvement of executive functions in prospective memory performance of adults. *International Journal of Psychology, 38*, 195-206.

Martin, J. S., Poirier, M., & Bowler, D. (2010). Brief Report: Impaired temporal reproduction performance in adults with autism spectrum disorder **.** *Journal of Autism and Developmental Disorders, 40,* 640-646, DOI: 10.1007/s10803-009-0904-3

McCartney, M., Burchinal, M.R., & Bub, K.L. (2006). Best practices for Quantitative methods for developmentalists. *Monographs of the Society for Research in Child Development*, *71*, 1-145

Nelson, H. E. (1976). Modified card sorting test sensitive to frontal lobe deficits. *Cortex, 12*(4), 313-324.

Obonsawin, M. C., Crawford, J. R., Page, J., Chalmers, P., Cochrane, R., & Low, G. (2002). Performance on tests of frontal lobe function reflect general intellectual ability. *Neuropsychologia, 40*(7), 970-977.

Passler, M.A., Isaac, W., & Hynd, G.W. (1985). Neuropsychological development of behavior attributed to frontal lobe functioning in children. *Developmental Neuropsychology, 1*, 349-370.

Skuse, D., Warrington, R., Bishop, D., Chowdhury, U., Lau, J., Mandy, W., et al. (2004). The developmental, dimensional and diagnostic interview (3di): A novel computerized assessment for autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*(5), 548-558.

Szelag, E., Kowalska, J., Galkowski, T., Poppel, E. (2004). Temporal processing deficits in children with high-functioning autism. *British Journal of Psychology, 95,* 269-282.

Tulving, E. (1985). How many memory systems are there? *American Psychologist, 40,* 385–398. doi:10.1037/0003-066X.40.4.385

Ward, H., Shum, D., McKinlay, L., Baker-Tweney, S., & Wallace, G. (2005). Development of prospective memory: Tasks based on the prefrontal-lobe model. *Child Neuropsychology, 11,* 527-549.

Wilson, B. A. C., & Baddeley, A. D. (1991). *The Rivermead Behavioural Memory Test* (2nd ed.). Bury St Edmunds, UK: Thames Valley Test Company.

Williams, D. (2010). Theory of own mind in autism: Evidence of a specific deficit in self-awareness? *Autism*, *14*, 474-494.

Williams, D.M., Bowler, D.M., & Jarrold, C. (2012). Inner speech is used to mediate short-term memory, but not planning, among intellectually high-functioning adults with autism spectrum disorder. *Development and Psychopathology 24*, 225-239.

Williams, D., & Happé, F. (2009). 'What did I say?' versus 'What did I think?': Attributing false beliefs to self amongst children with and without autism. *Journal of Autism and Developmental Disorders, 39*, 865-873.

Williams, D., & Happé, F. (2010). Representing intentions in self and other: Studies of autism and typical development. *Developmental Science*, *13*, 307-319.

World Health Organisation. (1993). *International classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines (10th edn.)* Geneva: World Heath Organisation.

Table 1: *Participant characteristics*

	Group		t	p	Cohen's *d*
	ASD (n = 21)	TD (n = 21)			
Age	10.60 (2.01)	10.59 (1.31)	0.02	.99	0.01
VMA	11.44 (3.33)	11.73 (2.69)	0.31	.76	0.10
VIQ	103.57 (17.88)	106.48 (14.01)	0.59	.56	0.18
PIQ	110.19 (16.35)	107.48 (13.23)	0.59	.56	0.18
SRS Raw Score	115.62 (24.34)	30.10 (20.63)	11.99	< .001	3.79

SRS: Social Responsiveness Scale

Table 2: Time-based and event-based PM task performance and overall ongoing task performance among ASD and TD participants (Means and SDs)

	Group		
	ASD	TD	*d*
Number of Time-based Failures	1.57 (1.36)	0.81 (0.81)	0.68
Number of Event-based Failures	0.29 (0.72)	0.43 (0.93)	0.17
Overall Ongoing Task Performance	312.02 (117.50)	367.26 (103.54)	0.50

Table 3: Mean group scores (and SDs) and group differences on the WCST and Animations task

	Group				
	ASD	TD	*t*	*p*	*d*
WCST Perseverative Errors	8.19 (5.76)	4.19 (3.43)	2.74	.009	0.84
WCST Non-perseverative Errors	5.86 (4.08)	5.19 (3.23)	0.59	.56	0.18
Animations Mentalising Score	0.55 (0.83)	1.81 (1.63)	3.09	.004	0.97
Animations Physical Score	3.15 (1.04)	2.76 (1.26)	1.07	.29	0.34

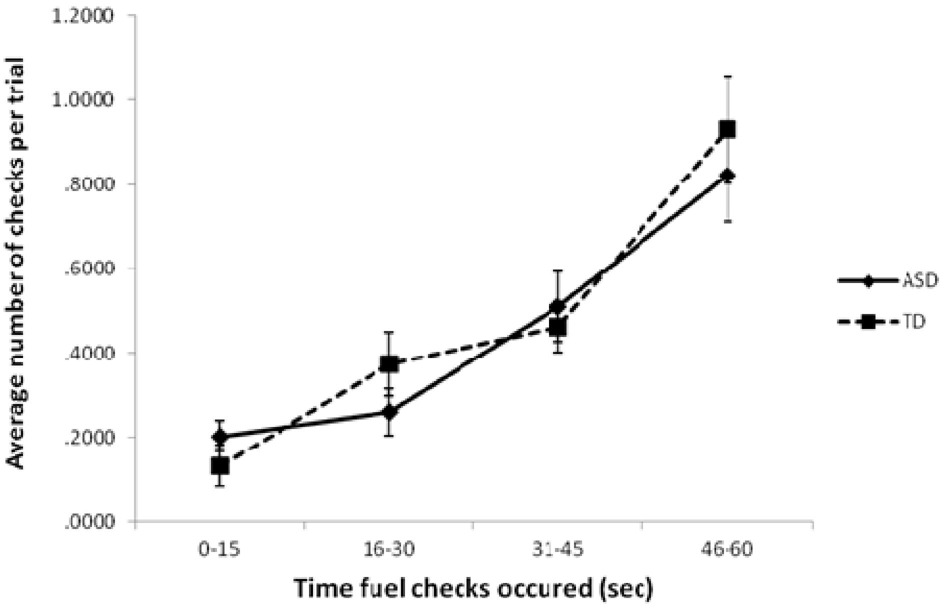
Table 4: Partial correlations between WCST and Animations task, and number of time-based PM failures and overall ongoing task performance among ASD and comparison participants

Measures	Group	Number of time-based PM failures	Overall ongoing task performance
WCST: No. of perseverative errors (controlling for No. of non-perseverative errors)	ASD	.23	-.29
	Comparison	.32	-.13
Animations task: Mentalising score controlling for Physical score	ASD	-.44*	.21
	Comparison	.13	-.30

*Correlations were statistically significant p < .05, 1-tailed

Figure 1: Illustration of the ongoing task

Figure 2: Average number of clock checks (averaged across trials) made by ASD and TD participants during the time-based PM task



Notes

1) In our view, groups should only be considered matched on a variable if group differences in that variable are associated with small/negligible effect sizes (i.e., $d \leq 0.50$) (see McCartney, Burchinal, & Bub, 2006; see also Mervis & Robinson, 2003).

2) Of course, in everyday life it may be that individuals with ASD display PM failures precisely because they need to devote more attention than is typical to other ongoing (day-to-day) activities in order to complete them. Thus, there may be a general difficulty in ASD with carrying out more than one task (cf. Mackinlay, Charman, & Karmiloff-Smith, 2006). The aim of this study, however, is to establish whether there is a specific problem with PM, and what the cognitive correlates of any specific impairment are.

3) Adjusted group means for number of time-based PM failures after controlling for ongoing task performance were, ASD: M = 1.50, SE = .24; Comparison: M = 0.89, SE = .24. Adjusted group means for number of event-based PM failures after controlling for ongoing task performance were, ASD: M = 0.20, SE = .17; Comparison: M = 0.51, SE = .17.

4) It should be made clear that the focus of Jones et al.'s (2011) paper was not on PM specifically, but rather on everyday memory problems in ASD. Of course, in everyday life, a failure of retrospective memory could lead directly to poor PM with all the negative consequences associated with diminished PM apparent as a result. Thus, there is no critique of Jones et al.'s approach implied in our discussion. Rather, we wish merely to understand whether PM is impaired in ASD, independent of diminished retrospective memory. Given this specific aim, we argue that our approach/analysis is more informative.

5) The accuracy of the following calculations has been kindly confirmed by C. Jones (personal communication, January, 2012). Taking the categorical data reported in Table 1 of Jones et al.'s paper, we

were able to calculate the group differences in rates of spontaneous PM success on each of the three key subtests. For the sake of parsimony, we report in full only the results from the Appointment subtest. According to our calculations, 78.0% (n = 39) of ASD participants and 88.6% (n = 39) of comparison participants spontaneously asked the experimenter the time when the alarm sounded. Twenty-two percent (n = 11) of participants with ASD and 11.4% (n = 5) comparison participants forgot to ask the experimenter despite recalling the instructions when asked. Therefore, once those individuals who failed this PM task because they had forgotten the instructions were excluded, the differences between the groups on this subtest was non-significant and associated with only a negligible/small effect size, $\chi^2 = 1.88$, $p = .17$, $\varphi = .14$. Using the same procedure, our calculations showed that group differences on the Message and the Belonging subtests were also non-significant and associated with negligible effect sizes (all ps > .38, all φs < .08).