J Autism Dev Disord (2011) 41:1445–1454
DOI 10.1007/s10803-010-1151-3

ORIGINAL PAPER

Investigating Multitasking in High-Functioning Adolescents with Autism Spectrum Disorders Using the Virtual Errands Task

Gnanathusharan Rajendran · Anna S. Law · Robert H. Logie · Marian van der Meulen · Diane Fraser · Martin Corley

Published online: 23 December 2010

Abstract Using a modified version of the Virtual Errands Task (VET; McGeorge et al. in Presence-Teleop Virtual Environ 10(4):375–383, 2001), we investigated the executive ability of multitasking in 18 high-functioning adolescents with ASD and 18 typically developing adolescents. The VET requires multitasking (Law et al. in Acta Psychol 122(1):27–44, 2006) because there is a limited amount of time in which to complete the errands. ANCOVA revealed that the ASD group completed fewer tasks, broke more rules and rigidly followed the task list in the order of presentation. Our findings suggest that executive problems of planning inflexibility, inhibition, as well as difficulties with prospective memory (remembering to carry out intentions) may lie behind multitasking difficulties in ASD.

Keywords Autism · ASD · Multitasking · Executive functions · Virtual environment

G. Rajendran (✉) · D. Fraser
School of Psychological Sciences and Health, University of Strathclyde, 40 George Street, Glasgow G1 1QE, Scotland, UK
e-mail: thusha.rajendran@strath.ac.uk

A. S. Law
School of Natural Sciences and Psychology, Liverpool John Moores University, Liverpool, England, UK

R. H. Logie · M. van der Meulen
Human Cognitive Neuroscience, Department of Psychology, University of Edinburgh, Edinburgh, Scotland, UK

M. Corley
School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, Scotland, UK

Introduction

In recent years, one of most influential cognitive theories of autism has been the executive dysfunction hypothesis. Despite a substantive body of research, however, findings have been mixed as to the existence and the characterisation of impairments in executive functions (EF) in autism spectrum disorders (ASD) (for reviews see Hill 2004; Hill and Bird 2006; Rajendran and Mitchell 2007). Some EF have been consistent in differentiating individuals with ASD from other groups; for example, planning (e.g. Tower of Hannoi, Tower of London tasks) seems to be poor in ASD (Hill 2004; Ozonoff and Jensen 1999). Whereas other EF, for example inhibition (e.g. Stroop), appears less successful at discriminating individuals with ASD from those without ASD (Hill 2004). Further, EF profiles may differ across neurodevelopmental disorders (Alloway et al. 2009; Ozonoff and Jensen 1999) and may offer a useful way of differentially characterising disorders.

In contrast to this sizeable literature, few (to our knowledge there are only four) studies have investigated *multitasking* in ASD—which could be argued to involve many EF. Multitasking has been operationalised as the ability to interleave tasks with one another; each being suspended and then resumed after appropriate intervals. Burgess (2000a, b; Shallice and Burgess 1991) has shown that multitasking is sensitive to the impact of frontal lobe damage in adults, even in cases where standard tests of EF are not. This suggests that multitasking paradigms might have a greater sensitivity for detecting EF difficulties, which would be beneficial when assessing clinical populations. Multitasking is necessary whenever time restricted tasks can neither be completed sequentially, nor simultaneously (Law et al. 2006; Logie et al. in press). A typical, everyday example of multitasking is the preparation of a meal—which involves numerous and varied sub-tasks, all of which have to reach completion at the same time (Craik and Bialystok 2006).

The first multitasking assessment in ASD was as part of the validation of a clinical test battery designed to measure executive dysfunction in children: the Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C, Emslie et al. 2003). Thirteen children with a Pervasive Development Disorder[1] (mean age = 9.2 years; mean IQ = 94), were compared with typically developing children and children with various clinical diagnoses (e.g. developmental co-ordination disorder, attention deficit disorder etc.). As part of one of the five BADS-C subtests, children performed the 'Six Parts Test'—inspired by the Six Elements Test (Shallice and Burgess 1991)—in which they had to attempt six tasks in 5 minutes with the order of task performance limited by rules. The ASD group had the lowest mean scaled scores of any group on this sub-test.

The second study, (Mackinlay et al. 2006) used a novel task called the Battersea Multitask Paradigm with fourteen high-functioning children with ASD (HF-ASD) and sixteen typically developing (TD) children (mean ages 12 years 0 months and 11 years 11 months respectively). The Battersea Multitask Paradigm was composed of three interleaved tasks (sorting, counting and colouring) that had to be completed in 3 minutes. This test generated six dependent measures based on the Greenwich Multitask Paradigm (Burgess et al. 2000), and these measures broke down performance into a six-stage invariant sequence (Rule Learn, Plan, Perform, Plan Follow, Monitor, Rule Memory). Mackinlay et al. found that the ASD group were significantly poorer in Plan and Perform, but there were no group differences in the four other stages (although the trend was in the direction of poorer performance in the HF-ASD group). The authors concluded, therefore, that it was in the areas of Planning and applying the rules (Performance) that underlay impaired multitasking performance in HF-ASD.

The third study, by Hill and Bird (2006) found that adults with Asperger syndrome performed more poorly than a typically developing adult comparison group on the Modified Six Elements sub-test of the adult version of the BADS (Wilson et al. 1996); specifically on the number of tasks completed and longest amount of time taken to complete a task.

The fourth, and most recent study on multitasking in ASD, was conducted by White et al. (2009) whose investigation included how children with ASD performed on the BADS-C. Mirroring Hill and Bird's (2006) finding in adults, White et al. found that children with ASD completed fewer tasks and took longer on any subtask than the TD control children in the 'Six Parts Test'. Notably, White and colleagues argued for the differentiation of EF tasks depending on whether they are "Constrained" or "Open-Ended". For example, the 'Playing Cards Task' (analogous to the Wisconsin Card Sorting Task, Heaton 1981) was considered to be "Constrained" because participants could not generate novel ways of doing the task—they simply had to follow a rule. In contrast, the 'Six Part Test' and the 'Key Search Test' were categorised as "Open-Ended"; in the case of the 'Key Search Test' this was because the participant had to line-draw an imaginary search pattern, but could do this in whatever way he or she wished, to find keys they had lost in a field. White et al. found that their ASD group generally performed more poorly on the "Open-Ended" in comparison to "Constrained" tasks than a TD group. White and colleagues went on to argue that problems with "Open-Ended" tasks in ASD might be due as much to the socio-communication difficulties in this group, as any EF problems per se. That is, "Open-Ended" tasks, by their very nature, are relatively unstructured and require the participant to infer what the experimenter wants from them (i.e. the participant is not *explicitly* told to do the task in the most efficient way). Using White et al.'s categorisation, it could be argued that multitasking tasks are intrinsically "Open-ended" and, therefore, performance on such tasks might be influenced by socio-communication as well executive ability.

Present Study

We sought to build on previous work by using a paradigm designed (1) to be high in ecological validity (Geurts et al. 2009), (2) be "Open-Ended" enough to generate different strategies and (3) provide multiple measures as is possible with other tests of multitasking (e.g. the Modified Six Elements Task and Battersea Multitask Paradigm) to illuminate the process of multitasking and so view how/why it might break-down. Accordingly, we used the Virtual Errands Test (VET; McGeorge et al. 2001) which is similar to the Virtual Multiple Errands Test (Rand et al. 2009a, b) in which real world-type errands had to be completed in a virtual environment. Multitasking in a virtual environment has not, hitherto, been used to test individuals with ASD.

McGeorge et al. (2001) designed the VET to compare five Dysexecutive patients and five matched controls virtual and real-life Multiple Errands Test performance in an office building. The tasks were a series of simple office-type errands (e.g. meet a colleague, send a fax from the main office, check details on a noticeboard etc.). Patients completed significantly fewer errands than controls in both

[1] It is not explicitly stated in the test manual, however, what kind of Pervasive Developmental Disorder the children had. Our assumption—like that of Mackinlay et al. (2006)—is that the children had an ASD.

real and virtual environments—a finding echoed by Rand et al. (2009a). Thereby suggesting that the VET was just as sensitive to dysexecutive impairment as the real-world version. Our study used essentially the same virtual environment as McGeorge et al. (2001), but with the office environment transformed into a school and with tasks appropriate for a school teacher. This was done to make the tasks as comprehensible as possible for the school children participants who were asked to run errands for their teacher.

We predicted that although participants with ASD would be able to do the tasks individually when presented singly and serially, that they would have problems when these tasks were presented simultaneously and had to be completed within a time limit; that is when they had to multitask. From the previous literature on ASD, we anticipate that these problems would be manifest in failures of on-line planning, possibly performing the tasks in an inefficient order, and breaking rules for moving around the virtual environment.

Method

Participants

Thirty-six adolescents took part in this study: 18 with ASD and 18 who were typically developing. The eighteen adolescents with ASD were recruited from either mainstream schools with a specialist autism unit, or a school for children with Special Educational Needs—all within Scotland. All of the participants with ASD had received an official diagnosis based on DSM-IV-TR criteria (American Psychiatric Association 2000) and had a Statement of Special Educational Needs. Additionally, the Social Communication Questionnaire (SCQ: Rutter et al. 2003), was completed by the parents of 10 of the 18 adolescents with ASD (55.5% completion rate); participants' mean SCQ score was 25.4, SD = 2.2.

All participants were tested with the Wechsler Abbreviated Scales of Intelligence (WASI; Wechsler 1999) which consists of four subtests: Two assess receptive Verbal Intelligence (VIQ) and two assess Performance Intelligence (PIQ)—sometimes referred to as non verbal IQ. Both scales combine to give a Full-Scale score (FSIQ). Table 1 provides details of gender, chronological age (CA), verbal IQ (VIQ), performance IQ (PIQ), full-scale IQ (FSIQ), and BADS profile and BADS Modified Six Elements Test scores (see below).

The groups did not significantly differ on gender: χ^2 (1, $N = 36$) = 0.8, $p = .37$; age: F (1, 34) < 1; VIQ: F (1, 34) < 1. However, the groups did differ on FSIQ: F (1, 34) = 7.6, $p < .01$; PIQ: F (1, 34) = 20.6, $p < .001$; and BADS profile (see below): F (1, 34) = 19.5, $p < .001$.

Materials

A Toshiba F10 Quosmio laptop was installed with the VET (McGeorge et al. 2001) which was built using Superscape 3D Webmaster and run using Superscape Visualiser. VET performance was recorded using screen capture software (Flashback by Blueberry Software). A manual stopwatch was used for timing.

The original design of the VET was based on an actual university building, and the task was to role-play a lecturer running a set of errands. The building consisted of three floors connected by two stairwells. Each floor was made up of a long corridor, with consecutively numbered offices along one side. Stairwells are reached through doors on the other side of the corridor, and participants were told only to

Table 1 Participants' characteristics

Participants	Age	Gender ratio	IQ			BADS profile score (max = 24)	BADS modified six element subtest (max = 4)
			Verbal	Performance	Full Scale		
ASD ($N = 18$)							
Range	11.6–17.4	16 M/2F	85–132	63–109	74–115	7–23	1–4
Mean	13.9		106.2	87.6*	96.2*	14.4*	2.5*[a]
(SD)	(1.7)		(14.6)	(14.8)	(13.1)	4.4	(1.2)
TD ($N = 18$)							
Range	12.2–18.3	14 M/4F	88–129	91–119	88–126	17–22	1–4
Mean	13.8		106.4	106.1	106.8	19.2	3.4
(SD)	(1.4)		(12.2)	(8.9)	(10.0)	1.5	(0.9)

ASD autism spectrum disorder, *TD* typically developing group

* $p = < .01$

*[a] $p = < .01$ using an ANCOVA covarying PIQ

travel in one direction for each staircase. In our version of the task the computer program was not altered, but the building was described as a School, and participants were put in the role of a pupil at the school and that their teacher had asked them to run some errands. None of the participants was familiar with the real building on which the VET was based. The tasks chosen for the errand list were a subset of the ones available in the computer program, and ones that might plausibly be conducted within a school. Examples are “Go to the door of F15 and click on the blue notice to find out the date of the exam” and “Collect a book from room S15 and deliver it to room S9”. There were two different errand lists, each with 6 errands. Three of the errands only had one “step” to them, as in the first example above. Two of the errands had two steps to them, as in the second example above, and one errand on each list had three steps. This meant that there were 10 possible subtasks to complete. There were 2 versions of the Errand List, which involved slightly different rooms and objects but had the same overall structure. Two versions of the task were included to see if performance generalised across both versions. So, if similar performance could be shown across both versions, then we could be more confident about the task’s validity. Exactly half of the participants in each group received List 1, and half received List 2.

A view of the environment was presented in a window in the centre of the screen, with a black frame around the side divided into boxes with grey lines (see Fig. 1). Boxes on the left displayed items that the participant had collected and boxes on the right displayed information that he or she had discovered. At the top, feedback was provided when they completed a task successfully (e.g. “Well done you have delivered the book”) or when they attempted to do something incorrectly. The top right corner displayed the number of the last door the participant “clicked on it” with the mouse, allowing participants to navigate their way around the building, in the same way as examining room numbers on doors in an actual building.

Pre-Assessment

In addition to the WASI test of intelligence, each participant was also tested with the Behavioural Assessment of the Dysexecutive Syndrome (BADS; Wilson et al. 1996), which, according to Evans et al. (1997), is an ecologically valid test of EF, assessing the everyday difficulties

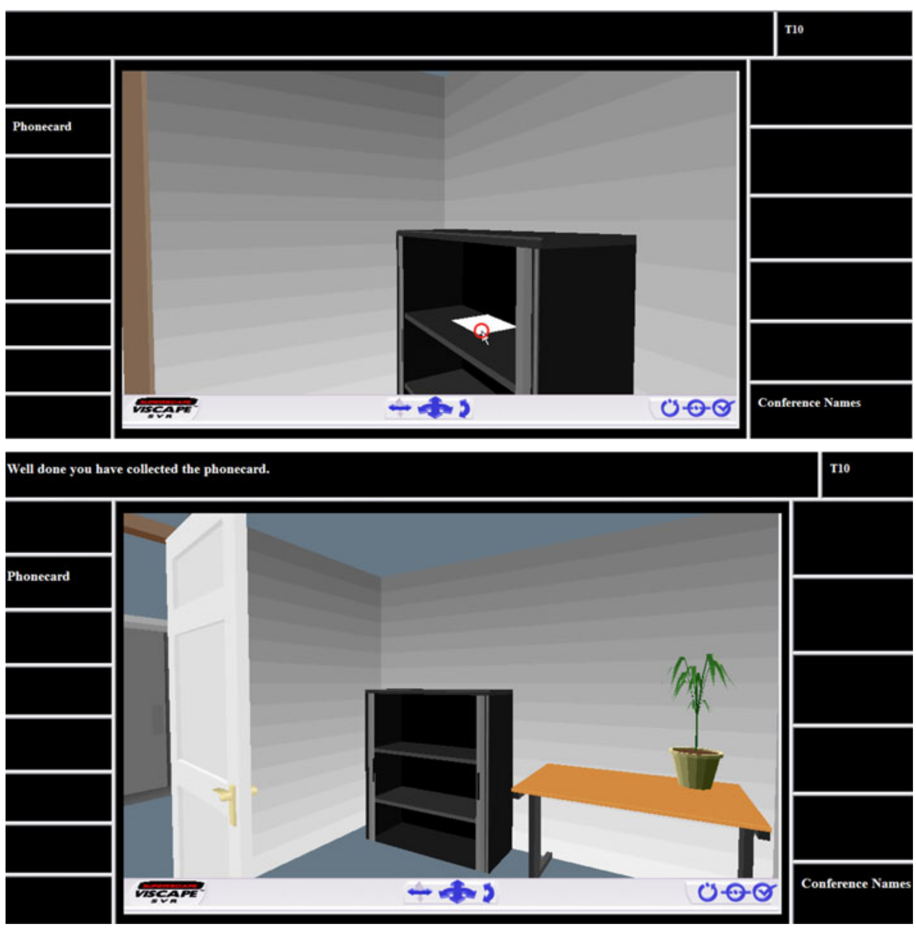

Fig. 1 Screenshot of the VET display

associated with dysexecutive syndrome. The BADS is composed of six subtests and assesses a range of cognitive functions representative of executive abilities. For example, the Modified Six Elements subtest purports to measure multitasking. For each subtest a summary profile is obtained (with a maximum of 4 and minimum of 0), and these are summed to produce an overall profile score out of 24. From the BADS handbook (Wilson et al. 1996), a score of between 16 and 20 is classified as average, and the mean score from the 78 brain injured patients, from whom the normative values were derived, is 14.03. (for more detailed descriptions of the BADS, its subtests, and the cognitive processes it purports to tap into, see Evans et al. 1997; Hill and Bird 2006; Norris and Tate 2000; Rajendran and Mitchell 2006).

The BADS is the original and adult version of the BADS-C (Emslie et al. 2003). The BADS-C has, in essence, the same subtests and structure as the BADS but was designed for 7–16 year olds. The original BADS was used in this reported study because the oldest participants were 17 and 18 and, so, too old for the BADS-C. In terms of appropriateness for the youngest participants; previous research has found it suitable for children as young as 11 years (Rajendran et al. 2005).

Procedure

All participants were tested in a quiet space either in their place of education, home, or (in the case of some of the TD participants) in the Psychology building, the University of Edinburgh. Most participants were tested for between three and four separate sessions; with each session lasting a maximum of 40–50 minutes.

The experimental session comprised the VET Training task, VET Screening Errands and the VET itself. In the Training task, instructions were read out to the participant who was given a map of the building to consult (see Fig. 2)—same school layout was used throughout the training, screening and task phases.

Participants were told that the map was a rough guide, and did not show every room in the environment, just the ones needed to complete the tasks. The purpose of the Training task was to familiarise participants with the mouse-based controls for moving around the environment. They were guided on a tour of the building that encompassed some key rooms, and were also shown how to complete a task (making a cup of hot chocolate). This was accomplished by collecting chocolate in one room, milk in another, and going to a third room where they clicked on a kettle to make the chocolate. During this training the participants were told about and asked if they understood that the room labels *F*, *S* and *T* referred to First, Second and Third floors respectively. Additionally, they were instructed

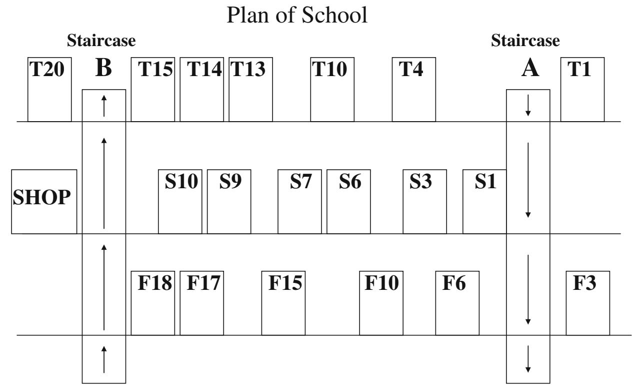

Fig. 2 Training Map

about the stairwell rule, which stated that they could only go up staircase B, and down staircase A.

The purpose of the Screening Errands, which followed the Training task, was to make sure that participants understood how to do individual tasks in the environment without any input from the experimenter. The environment was then re-loaded, but with new tasks, and the "start point" in a different room. Participants were first of all given a simple "one step" errand (collect a folder from room S7). When they successfully completed that, they were given another one step errand. This was followed by a two step and then a three step errand. All participants were able to complete these errands one at a time, and went on to the multitasking part of the session.

For the VET proper, participants were given a list of errands on a piece of paper, with the following instructions at the top.

"Imagine you are a student at the school. Your teacher is very busy and asks you to run some errands for her. These errands are listed below. She tells you that you can do the errands in any order you like, and that you can switch from one errand to another at any time. She also tells you that you must not go into any room except for those mentioned on the list. You have 8 minutes to complete as many of the errands as you can".

As soon as the participants had read through the sheet, a stopwatch was started and placed in front of the computer screen so that participants could see how much time they had remaining. They kept the sheet of paper with the list of errands and the rules with them throughout the task. The participants no longer had reference to a map of the building, given the thorough familiarisation procedure that had gone before.

When the 8 minutes of the task was up, participants were asked to stop what they were doing (unless they finished before the time limit). Their performance on the VET was recorded using the screen capture software, for later analysis.

Dependent Measures

Following Law et al. (2006), the main dependent measure on the VET task was "Score". To calculate this, participants were awarded a point for each part of an errand that was completed, but a point was deducted for any Action Error (e.g. picking up the wrong object) or Room Error (going into the wrong room). This general principle of awarding points for task completions and subtracting them for rule breaks has also been used with other tests of multitasking (e.g., Burgess et al. 2000; Mackinlay et al. 2006). The maximum possible score on the VET was 10, if a participant completed all the errands and made no errors. A count was also taken of the number of trips a participant made up or down one of the staircases. Given that using the stairwells was one of the most time-consuming activities in the virtual environment, a higher number of stairwell journeys was considered to reflect a less efficient strategy. A separate count was made of how often participants broke the rule governing the use of stairwells—i.e., only to go up Staircase B and down Staircase A. Further analyses of strategy looked at how much participants chose to re-order the errands given in the task list (see below) and how many 1, 2 and 3-step errands were completed by each group. If participants finished all the errands before the 8 minutes time limit, a note was taken of the finishing time.

Results

Fourteen of the 18 in the ASD group did not finish within the 8 minutes time limit, compared with 11 of the 18 in the TD group (χ^2 (1, $N = 36$) = 1.18, $p = .28$). Of those who did finish within time, the ASD group took a mean average 7:02 minutes and the TD group 6:55 minutes; a difference that was not significant (F [1, 9] = < 1).

Table 2 shows participants' performance on various VET measures. Given the IQ profile of the ASD group (a higher VIQ than their PIQ) and the significant difference between the groups on PIQ (and consequently FSIQ), ANCOVAs co-varying PIQ were conducted. An initial ANCOVA analysis showed no interaction between which of the two versions of the VET task had been used and group (F (1, 31) < 1); accordingly, the results below are collapsed across task version.

The overall mean VET Score for the TD group was 8.61 (SD = 1.2), while the overall mean for the ASD group was 5.61 (SD = 2.4). An ANCOVA revealed that the ASD group performed significantly worse on VET Score than the TD group: F (1, 33) = 6.23, $p = .02$, partial $\eta^2 = .16$, but also that PIQ accounted for a similar amount of variance in VET Score between the groups (F [1, 33] = 6.54, $p = .02$, partial $\eta^2 = .17$).

Performance can also be measured in terms of the efficiency of the route that participants take when travelling around the virtual environment. The more efficient the route, the fewer the number of journeys up and down the stairs. ANCOVA revealed that the ASD group made significantly more Stairwell Journeys than the TD group: F (1, 33) = 13.2, $p = < .01$, partial $\eta^2 = .29$, and also that PIQ did not account for any variance in this difference between the groups (F [1, 33] = < 1).

One important rule in the VET is to only travel in one direction on each staircase. Analysing the number of rule breaks as a proportion of total number of stairwell journeys made by each participant, the ASD group made significantly more Stairwell Rule Breaks than the TD group: F (1, 33) = 6.21, $p = < .01$, partial $\eta^2 = .16$. Control participants broke the rule an average of 7.3% of the time they used the stairs, while on average ASD participants broke the rule 41.6% of the time. Performance IQ made a non significant contribution to this difference (F [1, 33] = 1.86, $p = .18$). This finding is supported by categorical

Table 2 Participants' VET performance scores

Participants	Parts (max = 10)	Action errors	Room errors	VET score (max = 10)	No. of corridor journeys	No. of stairwell journeys	No. stairwell rule breaks
ASD (N = 18)							
Range	2–10	0–3	0–3	0–9	8–16	4–8	0–5
Mean	7.06	.78	.67	5.61*	12.9	5.94**	2.44**
(SD)	(2.3)	(1.1)	(0.8)	(2.4)	(2.3)	(1.2)	(1.8)
TD (N = 18)							
Range	7–10	0–2	0–0	5–10	10–20	3–6	0–2
Mean	9.00	.39	.00	8.61	12.7	4.22	0.33
(SD)	(1.0)	(0.7)	(0.0)	(1.2)	(2.6)	(1.0)	(0.6)

* $p = < .05$ using an ANCOVA covarying PIQ

** $p = < .01$ using an ANCOVA covarying PIQ analysis, revealing that participants in the ASD group were much more likely to break this rule than participants in the TD group. Specifically, 14 out of 18 ASD participants broke the rule at least once (the highest number was 5 rule breaks) whereas only 5 of the 18 TD participants travelled along the stairs in the wrong direction (4 participants broke the rule once, 1 broke it twice), χ^2 (1, $N = 36$) = 9.0, $p < 0.01$.

Given the high number stairwell rule breaks, the data were examined for other evidence that the ASD group were more likely to break task rules. In the instructions given to participants at the top of the errand list, they were told not to go into any rooms other than those stated. However, 9 of the ASD participants did go into other rooms (1 of these went into 3 rooms), while none of the TD participants showed this behaviour. Analysis revealed a significant association between group membership and the likelihood of entering prohibited rooms χ^2 (1, $N = 36$) = 12.0, $p < 0.01$.

A possible reason why ASD participants made more stairwell journeys might be that they adhered more closely to the order that tasks were presented on the list, rather than re-ordering them efficiently. A "list following" score was derived in order to assess this possibility. Participants were given a point if they completed a task in the same serial position as it appeared on the list (e.g. if the third task they completed was the third task on the list, this was awarded a point). For any errands that were not completed in the same serial position as on the list, participants were then given a point if a pair of tasks was in the same sequential order as the list. For example, they would be given a point if the final two tasks they completed were the first two tasks on the list, as long as they were completed in the same sequential order. The total number of points was then divided by the total number of tasks the participant had completed, to give a proportion score for list-following. A score of 1 would indicate perfect adherence to the list, and the lower the number the more they deviated from list order. Accordingly, we found that the ASD group had a mean List Following Score of 0.63 (SD = 0.32) compared with the TD group who had mean score of 0.35 (SD = 0.19). ANCOVA showed that this between group difference was significant: F (1, 33) = 7.05, $p < .05$, partial $\eta^2 = .18$ and also that PIQ did not account for any variance in this difference between the groups (F [1, 33] = < 1).

In terms of the number of 1/2/3 step tasks completed (three of the six errands only had one "step" to them. Two of the errands had two steps to them, and one errand on each list had three steps), we found that the ASD group completed 63% of 1 step, 56% of 2 step, and 72% of 3 step errands. In contrast, the TD group completed 85% of 1 step, 92% of 2 step, and 72% of 3 step errands. In a 2 × 2 mixed ANCOVA (covarying PIQ), with Group as a between subjects factor and Errand Type as a within subjects factor, the interaction approached significance (F [2, 66] = 3.06, $p = .053$).

Correlative Relationships Between VET Performance and the BADS

An ANCOVA on the Modified Six Elements sub-task of the BADS, which attests to measure multitasking, revealed that the ASD group (mean = 2.5, SD = 1.2, max. possible score = 4) performed significantly worse than the TD group (mean = 3.44, SD = .9): ANCOVA, F = (1, 33) = 9.5, p = < .001, partial $\eta^2 = .22$, and also that PIQ did not account for any variance in this difference between the groups (F [1, 33] = 2.42, $p = 0.13$). However, when correlating VET Score and Modified Six Elements, for both groups separately and partialling out PIQ, no significant correlations were found: ASD group ($r = .36$, $p = .16$), TD group ($r = -.27$, $p = .29$).

Discussion

As a group, adolescents with ASD showed difficulties with multitasking in everyday-type tasks in a virtual environment. That is, tasks within the VET during the training phase—when given singly—were achieved. However, when these tasks had to be interleaved during the task phase, then the ASD participants' difficulties with this executive ability were expressed; i.e., they completed fewer tasks and made more errors in the allocated time. Notably, using gross measures of one's ability to compete the task, like time to completion, did not distinguish the groups. So, arguably the more fine-grained measures, like VET Score or List Following Score, offer a more sensitive measure of multitasking in ASD.

It might be argued though that these multitasking difficulties simply reflect a relative weakness in PIQ because, despite being well matched on VIQ, the groups were not matched on PIQ. However, using a very conservative test—ANCOVA (see Miller and Chapman 2001, for a discussion)—which risked a Type II error, we still found a significant difference between the groups. The lack of PIQ matching is, nevertheless, a limitation and our findings must be viewed in this light; not least because PIQ made a significant contribution to the between group variance on the main measure of multitasking (VET Score). Ideally, the clinical and comparison group should be matched a priori; however the uneven profile and substantive difference between the ASD group's VIQ and PIQ resulted in *post hoc* statistical solution to try to account for these differences. With this caveat, we argue that in general terms our findings support those of others who have found multitasking difficulties in ASD (Emslie et al. 2003; Hill and Bird 2006; MacKinlay et al. 2006; White et al. 2009). More importantly, these findings have arisen from the use of a novel paradigm involving a virtual environment that couples simulation of a real environment with the scientific control of a laboratory setting for testing participants and collecting high quality data. Our ASD participants also performed more poorly on the Modified Six Elements Test sub-task of the BADS, a finding that provides some cross-validation of the novel VET task for assessing ASD. It is also a finding that is consistent with Emslie et al. (2003), Hill and Bird (2006) and White et al. (2009), whose participants all demonstrated difficulties with either adult or child versions of that test. Additionally, our ASD participants performed more poorly on a number of performance measures from the VET, and we discuss the possible reasons for this below.

Inflexibility of Planning Processes

In terms of following the task list, the ASD group seemed to adhere to the order on the list more than the TD group; thereby suggesting an inflexibility and rigidity in sticking to the list order even when it would be more efficient to cluster errands on the same floor and complete them at the same time. Added to this it appears that the ASD group did not find the 3 step tasks any more difficult than the TD group. Rather it was in the lack of completion of 2 step and 1 step tasks that seems to have brought down the group's mean VET Score. *Prima facie* this seems counter intuitive because the ASD group might be expected to have greater difficulty with the ostensibly more difficult longer step instructions. So, this result is insightful in that it was not the complexity of number of steps in a task per se that was problematic, but instead a more general lack of efficiency in the ordering of the tasks, leading to an increased frequency of time-consuming stairwell journeys. It is possible that impairments in executive functioning underlie ASD participants' inability to engage in on-line planning while performing the VET, forcing them to fall back on the strategy of completing the errands in list order. This would be consistent with the findings of Mackinlay et al. (2006) that children with ASD were poorer at planning a course of action for the Battersea Multitask Paradigm. Notably, while the map of the building was laid out in front of the participants during training and screening phases, it was not available for the VET proper. Hence, the absence of the map might be a mitigating factor in the planning problems seen in the ASD group.

However, a key question is whether the adolescents with ASD performed less well in the VET because of problems of EF per se, or because of something intrinsic to the task but unrelated to executive ability. Task demands appear to be crucial in understanding EF in ASD, and using White et al.'s (2009) classification of EF tasks as either being "Open-Ended'' or "Constrained"—multitasking tasks seem to fall under the umbrella of "Open-Ended". This is because these tasks, by their very nature, can be tackled using different strategies; strategies that are more likely (than "Constrained EF Tasks") to reveal difficulties with socio-communication and "autistic" idiosyncrasies. Importantly, any idiosyncratic reasons for multitasking problems in ASD may not be mutually exclusive to any executive reasons behind problems of multitasking (White et al. 2009). Nevertheless, the fact that participants are generally not *explicitly* told to do the tests of multitasking in the most efficient way cannot be ruled out as reason behind poorer performance in ASD groups. Ideally, the possible effects on performance of being explicitly told to do the task in the most efficient way, as opposed to relying on the implicit understanding that this should be the case, needs to be investigated.

Retrospective versus Prospective Memory Problems

Retrospective memory has been broadly defined as the "ability to remember previously learned information, facts, or events" (Cuttler and Graf 2009, p. 394) and prospective memory as the "ability to formulate intentions, plans and promises, to retain them, to recollect and carry them out at the appropriate time or in the appropriate context" (Cuttler and Graf 2007, p. 339). Further, prospective memory does not seem to be a unitary construct; for example a standardised self report, the PMQ (Prospective Memory Questionnaire: Hannon et al. 1995), is comprised of four subscales: long term episodic memory (e.g. 'I missed appointments I had scheduled'), short term habitual memory (e.g. 'I forgot to comb my hair this morning'), internally cued memory (e.g. 'I forgot what I came into a room to get') and the individual's ability to use prospective memory in aiding strategies ('I write myself reminder notes'). Similar distinctions are made in the Propsective and Retrospective Memory Questionnaire (PRMQ- Smith et al. 2000).

With respect to ASD, two recent studies have looked specifically at prospective memory. The first, Altgassen et al. (2009) found that 11 children with ASD performed more poorly than 11 TD children on a time-based prospective memory task; i.e. they had to remember to perform an activity—at particular times—in the midst of doing an ongoing activity. In the second study, Jones et al. (in press) looked at prospective memory in 94 adolescents with ASD using the Rivermead Behavioural Memory Test (RBMT: Wilson and Baddeley 1991, 2nd edn). Using three sub-tests measuring prospective memory ('Message', 'Appointment' and 'Belonging'), Jones and colleagues found poorer performance in the ASD group in 'Appointment' and 'Belonging' subtests. Notably, a cue from the environment was given for the tests: e.g. in the case of 'Belonging' the participant was cued to remind the experimenter about a hidden pen when the experimenter said ''We have finished the testing''.

In the current study, we found that the ASD group broke more stairwell rules than the TD group. A finding that is consistent with MacKinlay et al.'s (2006) study in which 7 out of 14 individuals with HF-ASD broke Performance rules (i.e. had problems in applying the rules) compared with 2 out of 16 in the TD group. So, our ASD group's difficulty with rules suggests that they had problems in either 'forgetting/not remembering' rules that had previous been learned (retrospective memory), or not 'bringing to mind' this rule while on task (prospective memory). We found that exactly half of the ASD participants (and none of the TD participants) disregarded the instruction not to enter task-unrelated rooms. Importantly, there was no requirement to remember this rule because it was stated on the sheet they had with them throughout the task. This suggests that the stairwell rule breaks might be part of a pattern of disorganised, rule-breaking behaviour rather—a general lack of inhibitory control—than due to specific retrospective memory failures. So, it seems unlikely that the ASD group had problems with retrospective memory, but they may have had difficulty in prospective memory, in 'bringing the stairwell rule to mind' at the crucial moment. Indeed, it may be that prospective memory not only involves memory, but it also links with planning and inhibition—and arguably this link seems to be elucidated in the VET.

In future research, participants could be asked if they remembered the stairwell rule, after having completed the VET—in order to better distinguish whether problems in remembering the stairwell rule were due to difficulties with retrospective, or prospective memory. Further, the VET offers a potentially good test (and training) environment for the role of cues in prospective memory in ASD in everyday-type tasks. For example, we could investigate whether adolescents with ASD make stairwell rule break errors if environmental retrieval cues are provided for this rule while they are doing the VET.

Interestingly the Stairwell Rule Break is an entirely arbitrary rule and theoreticians have argued that individuals with ASD seem to have particular problems with rules in general, but especially arbitrary ones. For example, Cognitive Complexity and Control theory (CCC: Frye et al. 1995; Zelazo and Frye 1997) states that EF is related to theory of mind because both tests of theory of mind (e.g. the unexpected transfer task) and measures of executive ability involve higher order rule use. While Russell (Biro and Russell 2001; Russell 1997) argues that it is the arbitrary nature of the rules, which often occur in EF tasks, that individuals with ASD have particular difficulties with—possibly due to a lack of verbal self prompting (White et al. 2009). Our results suggest that problems with arbitrary rules in ASD may be due to problems of prospective memory: i.e. difficulties in remembering the intention to carry out a rule at the crucial moment.

Summary and Future Directions

For the first time, multitasking difficulties in adolescents with ASD have been found in everyday-type tasks, using a virtual environment. We argue that virtual environments offer all the necessary ingredients of single user agency in a real world type task, but with the benefits of ethical and experimental control. Additionally, using this method we can look at more sensitive, fine-grained measures than is possible in a genuine real-world environment or with more artificial laboratory tasks, to try and shed light on the possible underlying reasons behind any multitasking difficulties. Our research, therefore, has generated new findings with a novel paradigm, but also serves as a starting point for further investigation into the causes of multitasking impairments associated with ASD (e.g. planning, prospective memory, retrospective memory, time pressure anxiety, inhibitory control, socio-communication etc.). Finally, the VET offers the possibility of ameliorating difficulties with multitasking in ASD: either by investigating any influence of changing the task demands or environment on performance, and/or through training/teaching.

Acknowledgments This research was supported by a Research and Development Fund grant from the University of Strathclyde and the University of Edinburgh's Development Trust Research Fund to Gnanathusharan Rajendran. The authors wish to thank all the children, parents and schools with their help with this research. Multitasking research by Robert Logie and Anna Law is supported by Grant number F/00 158/BW from the Leverhulme Trust.

References

Alloway, T. P., Rajendran, G., & Achibald, L. M. D. (2009). Working memory in children with developmental disorders. *Journal of Learning Disabilities, 42*(4), 372–382.

Altgassen, M., Williams, T. I., Bolte, S., & Kliegel, M. (2009). Time-based prospective memory in children with autism spectrum disorder. *Brain Impairment, 10*(1), 52–58.

APA. (2000). *Diagnostic and statistical manual of mental disorders* (4th text revised ed.). Washington, DC: American Psychiatric Association.

Biro, S., & Russell, J. (2001). The execution of arbitrary procedures by children with autism. *Development and Psychopathology, 13*(1), 97–110.

Burgess, P. W. (2000a). Real-world multitasking from a cognitive neuroscience perspective. *Attention and Performance, 18*, 465–472.

Burgess, P. W. (2000b). Strategy application disorder: The role of the frontal lobes in human multitasking. *Psychological Research-Psychologische Forschung, 63*(3–4), 279–288.

Burgess, P. W., Veitch, E., Costello, A. D., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia, 38*(6), 848–863.

Craik, F. I. M., & Bialystok, E. (2006). Planning and task management in older adults: Cooking breakfast. *Memory and Cognition, 34*(6), 1236–1249.

Cuttler, C., & Graf, P. (2007). Sub-clinical compulsive checkers' prospective memory is impaired. *Journal of Anxiety Disorders, 21*(3), 338–352.

Cuttler, C., & Graf, P. (2009). Checking-in on the memory deficit and meta-memory deficit theories of compulsive checking. *Clinical Psychology Review, 29*(5), 393–409.

Emslie, H., Wilson, F. C., Burden, V., Nimmo-Smith, I., & Wilson, B. A. (2003). *Behavioural assessment of dysexecutive syndrome for children (BADS-C)*. Oxford: Pearson.

Evans, J. J., Chua, S. E., McKenna, P. J., & Wilson, B. A. (1997). Assessment of the dysexecutive syndrome in schizophrenia. *Psychological Medicine, 27*(3), 635–646.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10*(4), 483–527.

Geurts, H. M., Corbett, B., & Solomon, M. (2009). The paradox of cognitive flexibility in autism. *Trends in Cognitive Sciences, 13*(2), 74–82.

Hannon, R., Adams, P., Harrington, S., FriesDias, C., & Gipson, M. T. (1995). Effects of brain injury and age on prospective memory self-rating and performance. *Rehabilitation Psychology, 40*(4), 289–298.

Heaton, R. K. (1981). *Wisconsin card sorting test manual*. Odessa, FL: Psychology Assessment Resources.

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

Hill, E. L., & Bird, C. A. (2006). Executive processes in Asperger syndrome: Patterns of performance in a multiple case series. *Neuropsychologia, 44*(14), 2822–2835.

Jones, C. R. G., Happé, F., Pickles, A., Marsden, A. J. S., Tregay, J., Baird, G., et al. (in press). 'Everyday Memory' Impairments in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*. doi:10.1007/s10803-010-1067-y

Law, A. S., Logie, R. H., & Pearson, D. G. (2006). The impact of secondary tasks on multitasking in a virtual environment. *Acta Psychologica, 122*(1), 27–44.

Logie, R. H., Law, A. S., Trawley, S., & Nissan, J. (2010). Multitasking, working memory and remembering intentions. *Psychologica Belgica*, *50*, 309–326.

Mackinlay, R., Charman, T., & Karmiloff-Smith, A. (2006). High functioning children with autism spectrum disorder: A novel test of multitasking. *Brain and Cognition, 61*(1), 14–24.

McGeorge, P., Phillips, L. H., Crawford, J. R., Garden, S. E., Della Sala, S., Milne, A. B., et al. (2001). Using virtual environments in the assessment of executive dysfunction. *Presence-Teleoperators and Virtual Environments, 10*(4), 375–383.

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology, 110*(1), 40–48.

Norris, G., & Tate, R. L. (2000). The behavioural assessment of the dysexecutive syndrome (BADS): Ecological, concurrent and construct validity. *Neuropsychological Rehabilitation, 10*(1), 33–45.

Ozonoff, S., & Jensen, J. (1999). Brief report: Specific executive function profiles in three neurodevelopmental disorders. *Journal of Autism and Developmental Disorders, 29*(2), 171–177.

Rajendran, G., & Mitchell, P. (2006). Text chat a tool for referential questioning in Asperger syndrome. *Journal of Speech Language and Hearing Research, 49*, 102–112.

Rajendran, G., & Mitchell, P. (2007). Cognitive theories of autism. *Developmental Review, 27*(2), 224–260.

Rajendran, G., Mitchell, P., & Rickards, H. (2005). How do individuals with Asperger syndrome respond to nonliteral language and inappropriate requests in computer-mediated communication? *Journal of Autism and Developmental Disorders, 35*, 187–207.

Rand, D., Rukan, S. B. A., Weiss, P. L., & Katz, N. (2009a). Validation of the Virtual MET as an assessment tool for executive functions. *Neuropsychological Rehabilitation, 19*(4), 583–602.

Rand, D., Weiss, P. L., & Katz, N. (2009b). Training multitasking in a virtual supermarket: A novel intervention after stroke (Article). *American Journal of Occupational Therapy, 63*(5), 535–542.

Russell, J. (1997). How executive disorders can bring about an inadequate ''theory of mind''. In J. Russell (Ed.), *Autism as an executive disorder*. Oxford: Oxford University Press.

Rutter, M., Bailey, A., & Lord, C. (2003). *Social Communication Questionnaire*. Los Angeles, CA: Western Psychological Services.

Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal-lobe damage in man. *Brain, 114*, 727–741.

Smith, G., Della Sala, S., Logie, R. H., & Maylor, E. A. (2000). Prospective and retrospective memory in normal ageing and dementia: A questionnaire study. *Memory, 8*(5), 311–321.

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence (WASI)*. San Antonio: The Psychological Corporation.

White, S. J., Burgess, P. W., & Hill, E. L. (2009). Impairments on "Open-Ended" executive function tests in autism. *Autism Research, 2*(3), 138–147.

Wilson, B. A., Alderman, N., Burgess, P. W., Emslie, H. C., & Evans, J. J. (1996). *The behavioural assessment of the dysexecutive syndrome*. Oxford: Harcourt Assessment.

Wilson, B. A. C., & Baddeley, A. D. (1991). *The rivermead behavioural memory test* (2nd ed.). Bury St Edmunds, UK: Thames Valley Test Company.

Zelazo, P. D., & Frye, D. (1997). Cognitive complexity and control: A theory of the development of deliberate reasoning and intentional action. In M. Stamenov (Ed.), *Language structure, discourse, and the access to consciousness*. Amsterdam & Philadelphia: John Benjamins.