Journal of Autism and Developmental Disorders, Vol. 34, No. 4, August 2004 (© 2004)

The Use and Understanding of Virtual Environments by Adolescents with Autistic Spectrum Disorders

Sarah Parsons,[1,2] Peter Mitchell,[1] and Anne Leonard[1]

The potential of virtual environments for teaching people with autism has been positively promoted in recent years. The present study aimed to systematically investigate this potential with 12 participants with autistic spectrum disorders (ASDs), each individually matched with comparison participants according to either verbal IQ or performance IQ, as well as gender and chronological age. Participants practised using a desktop 'training' virtual environment, before completing a number of tasks in a virtual café. We examined time spent completing tasks, errors made, basic understanding of the representational quality of virtual environments and the social appropriateness of performance. The use of the environments by the participants with ASDs was on a par with their PIQ-matched counterparts, and the majority of the group seemed to have a basic understanding of the virtual environment as a representation of reality. However, some participants in the ASD group were significantly more likely to be judged as bumping into, or walking between, other people in the virtual scene, compared to their paired matches. This tendency could not be explained by executive dysfunction or a general motor difficulty. This might be a sign that understanding personal space is impaired in autism. Virtual environments might offer a useful tool for social skills training, and this would be a valuable topic for future research.

KEY WORDS: Virtual environments; social skills; adolescents; autistic spectrum disorder; executive function.

INTRODUCTION

The potential benefits of computer-based tasks for people with autistic spectrum disorders (ASDs) have been recognised in various recent articles (e.g., Moore, 1998; Moore, McGrath, & Thorpe, 2000; Moore & Taylor, 2000). In particular, computers offer a predictable and consistent environment in which the pace of working can be suited to individual needs (Swettenham, 1996). Visual and auditory inputs can be planned and controlled directly, allowing tasks to be completed without distraction from extraneous cues (Wilson, Foreman, & Stanton, 1998), and tasks can be presented consistently and repeatedly, without the boredom and fatigue that can sometimes occur with task repetition by human instructors (Cromby, Standen, & Brown, 1996). Interactive multimedia computer programs have been effective in teaching vocabulary (Heimann, Nelson, Tjus, & Gilberg, 1995) as well as improving motivation and attention during learning (Chen & Bernard-Opitz, 1993; Moore & Calvert, 2000). People with autism have also successfully learned about emotions (Silver & Oakes, 2001) and social problem solving (Bernard-Opitz *et al*., 2001) using computer-based tasks.

Virtual environments (VEs) are a specific kind of computer-based task, defined as a "…3D data set describing an environment based on real-world or abstract objects and data" (Blade & Padgett,

[1] School of Psychology, University of Nottingham, University Park, Nottingham, NG7 2RD, U.K.

[2] Correspondence should be addressed to Sarah Parsons, School of Education, University of Birmingham, Edgbaston, Birmingham, B15 2TT, U.K.; Tel.: +44-121-414-4819; Fax: +44-121-414-4865; e-mail: s.j.parsons@bham.ac.uk

2001; p. 26). This definition is based on what the user sees on the screen, rather than how the environment is encountered (e.g., through immersive headsets, or via a standard PC—see below). In VEs, realistic 3D scenes can be encountered in 'real time', which means that as a user navigates through a scene, the movement on the screen is like walking through a real environment at a normal pace. In the environments used for the present study, the user has a 'through-the-eyes' view of the scene, rather than a bird's-eye-view. This means that the user cannot see a representation of themselves on the screen, nor any part of themselves during task completion, such as hands and arms. VEs can include the representation of people as well as objects and have been used in a variety of fields for cognitive rehabilitation. For example, Rothbaum and Hodges (1999) have successfully used VEs to help attenuate vertigo and flying phobia in some of their patients, and Brown, Neale, Cobb and Reynolds (1999) have used VEs to help people with learning disabilities develop everyday skills.

Over and above the general benefits of being computer-based, VEs offer great potential for people with autism. Perhaps the most important advantage is that users can role-play in an environment designed to mimic specific social situations. The growing sophistication of VEs means that tasks and skills can be practised in increasingly realistic settings. This offers the potential for improved generalization of social skills between training and real-world environments (Strickland, Marcus, Mesibov & Hogan, 1996), as well as the possibility of encouraging mental simulation of events, which may aid social 'problem solving' (Parsons & Mitchell, 2002).

Consequently, there has been much optimism about the possible usefulness of VEs for members of the autistic population (Clancy, 1996; Parsons *et al.*, 2000; Trepagnier, 1999). However, there is a paucity of direct evidence to support the idea, with only a few published papers to date investigating this possibility. Strickland (1996), (Strickland *et al.*, 1996; see also Strickland's website www.do2learn.com for updated information about current research) used a fully-immersive system with two minimally verbal children with autism to see whether they would track the movement of cars down a virtual street by turning their heads, and 'walk' towards a street sign in the virtual scene, by walking a short distance in the testing room. Both children responded to the VE, at least on a basic level, and one child responded with her own physical movements in relation to the virtual scene. However, conclusions from this study are limited due to the small number of participants included, as well as some concerns about the applicability of fully-immersive systems. In particular, head-mounted displays can be very expensive, are heavy and may cause people to experience symptoms of 'cybersickness', such as nausea, headaches and dizziness (Cobb, Nichols, Ramsey, & Wilson, 1999). Fully-immersive systems may not currently be the most appropriate technology for use with some people with autism.

Desktop VEs represent an alternative approach, are presented on standard PC's and work with the aid of a joystick and a mouse. Desktop VEs are more affordable and accessible for educational use, and tend to be much less susceptible to the symptoms of cybersickness (Nichols, 1999). Eynon (1997) included some children with ASDs in his study and found that a desktop interface encouraged children to be focused and attentive whilst using the program. However, again conclusions are limited due to the lack of information regarding diagnosis and background ability, the lack of a comparison group and the small number of participants.

The present study was designed to systematically investigate the potential of VEs with a clearly specified group of participants with ASDs, alongside two carefully matched comparison groups. The main aim was to consider the two broad questions of use and understanding of desktop VEs. In terms of the use of VEs, we were interested in basic navigation skills, such as the time taken to complete different tasks, and the ease with which tasks were completed. The investigation of understanding VEs was based on the premise that VEs could support social skills training only if they were interpreted as having a representational quality. That is, participants need to understand that certain aspects of behavior within the environment can be achieved only by representing them in some other way. For example, eating and drinking are achieved by clicking on items with the mouse, and paying for items involves clicking on coins shown on the screen.

It is well known that people with ASDs often interpret situations and language literally, seemingly unaware of underlying meaning (e.g., Mitchell, Saltmarsh, & Russell, 1997). An overly literal interpretation of a virtual environment could limit its usefulness as a social skills teaching tool, as participants need to understand that what happens on the screen is not a literal representation of what happens in real life. Some evidence suggests that people with autism have difficulty understanding some forms of representation (e.g., beliefs; Leekam & Perner, 1991), but not others (e.g., pictures; Charman & Baron-Cohen, 1992). We reasoned that if participants understood VEs as representational, at least at the basic level of not actually being reality, then they would be willing to make a nonliteral interpretation about the behavior of some people in a virtual scene. To expand, participants were shown a Virtual Café (Brown *et al.*, 1999) in which two people (that is, virtual representations of people) are seen standing by the bar, facing each other, but not talking. However, their proximity and orientation towards each other suggests that they might be seen as engaged in conversation. If participants said that the people were talking to each other, this would suggest that they understood the figures in the scene to be representing interpersonal interaction. In addition, participants might answer in a similar way about a video of real people, filmed in a real bar. A video representation is different from virtual reality in the sense that it is a photo-realistic representation, which is familiar to people through the use of television and other videos. A virtual environment, by contrast, is less familiar to many people and is a more abstract form of representation. Comparing these representations enabled us to see whether participants interpreted them in similar ways.

A further area of investigation for the present study, was based on some observations made of a group of adult participants with Asperger Syndrome in a pilot study. Some members of the group showed a tendency to navigate between, or very near to, the couple standing at the bar in Virtual Café, en route to buy a drink. This behavior seemed to be socially inappropriate, especially as there was a large empty space at the bar, to the right of the couple. We were keen to investigate systematically whether this behavior would be shown by any of our participants in the present study and, if so, whether this effect would be more likely to be seen in the ASD group in particular. Of course, it could be that participants with ASDs have a difficulty navigating through virtual environments in a way that inadvertently leads to violations of personal space. Therefore, it was important to include general measures of navigational abilities.

Finally, we wanted to explore whether there were any links between use of VEs and performance on executive function tasks which assess a number of cognitive abilities such as inhibition of response, set-shifting and maintenance, planning and working memory (Pennington & Ozonoff, 1996). Much evidence suggests that people with ASDs show deficits on tests of executive function (e.g., Hughes & Russell, 1993; Hughes, Russell, & Robbins, 1994; Ozonoff, 1995), especially cognitive flexibility (Ozonoff, Strayer, McMahon, & Filloux, 1994), which has also been linked to social understanding (Berger *et al.*, 1993). Since the Virtual Café represents a social situation, it is possible that executive abilities may be involved in negotiating the environment. In addition, aspects of planning to achieve particular goals and inhibition of responses were required for successfully completing virtual training environments included in the present study, as well as in the Virtual Café. It may be that executive ability mediates the performance of people with ASDs on virtual environments, potentially providing information about where extra support may need to be incorporated within future designs of environments.

METHOD

Participants

There were 36 participants in total. Twelve were individuals diagnosed with an ASD (based on their school statement of special educational needs[3]), aged 13–18 years (see Table I for a full breakdown of background characteristics and relevant VIQ- and PIQ-matches). All demonstrated a full-scale IQ (FSIQ) of 70 or greater (assessed by the Wechsler Abbreviated Scale of Intelligence; Wechsler, 1999) except for one individual with a score of 66. According to DSM-IV (1994), full-scale IQ's above 70 qualify as 'nonretarded'. Each of the participants with an ASD was then individually matched with two other pupils from different schools: one matched on verbal IQ (VIQ) and the other matched on performance IQ (PIQ; both were also matched according to age and gender). The majority of participants who were matched according to VIQ were drawn from Special Needs Schools, and the majority of pupils

[3] All students attended a specialist school for students with ASDs and had received a diagnosis of ASD, or were described as having the autistic triad of impairments in communication, socialisation and imagination. The authors recognise the limitations of using school records to ascertain a specific diagnosis according to current diagnostic criteria, however no claims are being made about how diagnosis may be related to performance.

Table I. Individual Background Characteristics, and Mean Values, of the ASD Group and their Comparison Participants, Matched According to VIQ, PIQ and CA

	ASD				VIQ-match		PIQ-match	
ASD group member	CA	VIQ	PIQ	FSIQ	CA	VIQ	CA	PIQ
1. Male	13:3	55	97	75	12:7	59	13:3	99
2. Male	13:9	55	89	71	12:11	57	13:3	98
3. Male	13:11	55	87	70	12:10	64	14:0	79
4. Male	13:11	64	74	73	14:4	65	13:5	78
5. Male	14:3	55	77	72	13:6	57	14:3	72
6. Female	14:5	84	103	92	16:1	87	14:4	101
7. Female	14:7	110	86	98	14:7	110	14:8	84
8. Male	14:11	84	104	93	13:7	82	15:0	106
9. Male	17:1	75	96	84	16:6	71	16:8	97
10. Male	17:6	58	77	66	18:5	60	17:1	63
11. Male	18:1	55	101	76	17:11	55	16:10	103
12. Male	18:3	77	112	93	18:11	77	16:10	125
Mean	15:4	68.9	91.9	80.25	15:2	70.3	14:11	92.1
Range	13:3–18:3	55–110	74–112	66–98	12:7–18:11	55–110	13:3–17:1	63–125

matched for PIQ were drawn from a mainstream secondary school. Participants drawn from the latter school were matched according to scores obtained on a cognitive abilities test (CAT; NFER), commissioned by the school, 3 months prior to the present study. To avoid unnecessary 'over testing' we used the CAT scores already available, rather than test each PIQ-match pupil on the WASI. The CAT provides verbal, nonverbal and full IQ scores, and is standardised on a large sample of children across the UK (16,000). On their website, NFER (www.nfer-nelson.co.uk) consider the verbal and nonverbal batteries of the CAT to be analogous to the Wechsler Intelligence Scale for Children (WISC). Paired sample *t*-tests confirmed that there were no significant differences between the groups on the matching criteria. Permission for each person to take part was obtained from parents/guardians prior to participation.

Materials

All virtual environments were built in Superscape VRT and run on a laptop computer. Movement around the environments was achieved with a USB joystick and interaction with objects was achieved with the mouse. All sessions were digitally videotaped for later analysis. A quad-mixer was used to produce split-screen filming of participants: one section showed output directly from the computer screen, and another showed output from the video camera which was focused on the participant using the environments. Instructions and checklists were printed in large font, on white A4 card, and laminated. Pictures of objects used for the training sessions were printed directly from each of the different environments. The pictures were cut out individually, laminated, and mounted on background cards using Velcro.

Procedure

Tests of Executive Function

Prior to using the virtual environments, all participants completed five out of the six tasks[4] included in the Behavioral Assessment of the Dysexecutive Syndrome (BADS; Wilson, Alderman, Burgess, Emslie & Evans, 1986). This session was completed on a separate day prior to use of the virtual environments. The BADS is a battery of six standardised sub-tests, which provides a profile of executive function abilities. Rather than focusing on component executive function skills, each test within the battery taps a number of different aspects of executive functions, making the tasks more akin to solving everyday problems, which frequently require the co-ordination of a number of related abilities. A brief description of each task is included in the appendix, and full details can be found in Wilson *et al.* (1986).

[4] The 'Modified Six Elements Test' section of the BADS was not administered to the participants in this study. After starting to use it with some of the most high-functioning participants, it became apparent that the test was very difficult to complete.

The 'Training' Virtual Environments

Description of the Environment

These environments served a dual purpose. First, they were designed to give participants opportunity to practice using the equipment before entering the Virtual Café (see next section). We wanted to make sure that participants were comfortable with the joystick, could interact with objects on the screen by clicking on the mouse and had received the same exposure to our specific environments prior to taking part in the study. Second, they enabled us to investigate how well participants responded to different task demands in the VEs, such as navigating through open and confined space, completing a search task, and check whether performance changed over the course of four trials.

The training environment consisted of a simple building surrounded by open space (see Fig. 1(a)). Inside the building, there were maze-like corridors (see Fig. 1(b)). Moving around the outside and inside of the building enabled participants to practice maneuvering in open and confined spaces, respectively.

A large room was at the end of the corridor, which contained eight different objects. The objects could be seen only as a green square from a distance, but were revealed automatically as the user moved closer (Fig. 1(c)). Clicking on an object with the mouse moved the item to the black display bar at the bottom of the screen (see Fig. 1(c) in which the kettle has already been selected). There were five training environments in total. These were identical except for the objects in the room at the end of the corridor, which were categorised according to type of object: cars, kitchen utensils, electrical equipment, chairs, and tools.

Design and Procedure

Before participants used the equipment, the experimenter demonstrated the task and explained what participants would see and be expected to do. The experimenter showed participants a laminated sheet containing eight individual pictures of each of the objects in the first room (vehicles), and placed four of them on a laminated board with the numbers 1–4 printed along the top. The experimenter explained that the four pictures chosen would be the objects they would look for in a room at the end of a corridor, in the order they appeared on the card. The pictures chosen during this demonstration, and the order in which they appeared on the board, were identical for each participant. This board remained in clear view throughout the procedure.

Next, participants were shown how to navigate around the outside of the building and through the corridor of the building. Once inside the room at the end of the corridor, participants were shown that objects first appeared as green squares and could only be seen properly once the participant moved closer to them. The experimenter started to search the room to look for the first object shown on the numbered card, making sure that some

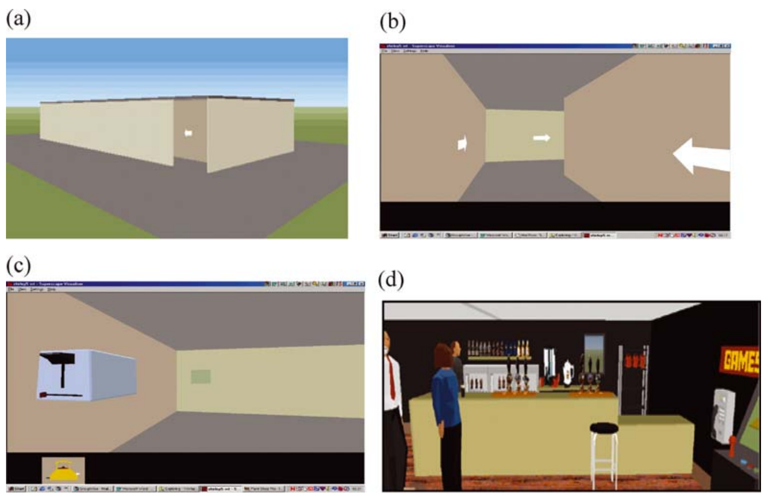

Fig. 1. (a) Outside view of the training environment; (b) corridor in the training environment; (c) near and distant views of objects in the training environment; (d) scene from the Virtual Café.

nontarget objects were revealed first. Participants were shown that clicking on an object with the mouse moved it from the room, down onto the display bar, and a further click moved it back into the room again (in the event of correcting any mistakes). After selecting the correct four objects, the experimenter noted that the display bar and the numbered card with pictures looked the same and emphasized that the task was "to make that part of the screen [point to display bar] look the same as the card [point to card]."

After watching the experimenter demonstrate the task, the participants completed four trials in the training environment. The training environments were identical apart from the different objects in the room at the end of the corridor. Each trial began with participants navigating around the outside of the building, and then entering the corridor in order to reach the 'object room'. Every participant completed the trials in the same order: kitchen utensils, electrical equipment, chairs, and tools. For each trial, the four objects to be found in the room were chosen at random, either by the participant or the experimenter. The person choosing the objects was alternated across the four trials, and whether the experimenter or the student chose first was counterbalanced between participants.

The Virtual Café

Description of the Environment

After completing the training trials, participants were given an opportunity to use the Virtual Café, an environment taken from the Virtual City of the Life Skills Project (Brown *et al.*, 1999). The environment was developed to support the learning of daily skills in people with general learning disabilities. It is more complex than the training environment in a number of ways. First, it looks more 'realistic' through the use of sophisticated textures and images. Second, there are people shown in the environment: a barman behind the counter, and two people standing facing each other at the bar (see Fig. 1(d)). Third, there are both textual and verbal prompts at different points within the café. For example, when the user moves close to a table they hear a voice in the program say 'would you like to sit here?', and two text boxes appear on the screen showing the words 'yes' and 'no'. The user needs to click on the text box to answer the question, before being able to progress. Finally, there are 'areas of interaction' in the café which flash red in order to prompt the user for a response. For example, there is a menu on the table that flashes red when the user sits down. Once the user clicks on the menu, the food and drink options appear on the screen, from which the user makes a choice.

Overall, both the training and the café environments required users to navigate with the joystick and interact with objects using the mouse. However, none of the more complex features of the café were included in the training environment. We were particularly interested in whether participants would experience problems completing the tasks in the café when complexity was added.

Design and Procedure

After completing the four training trials, participants were told that they were going to see a different environment and that there were different tasks to be completed. Participants were shown a checklist of tasks, and were encouraged to read each task aloud. They were told to complete the tasks in the order they were shown on the list. The checklist was visible throughout the session and listed the following tasks: (1) sit at the black table; (2) order some food and drink from the menu; (3) pay for the food and drink; (4) order a drink from the counter; (5) pay for the drink. Each task involved a number of subsidiary stages, which can be split into three different types:

1. Communication—the computer asks the user a question.
2. Interaction—the participant uses the mouse to interact with a part of the screen (e.g., responding to a question, or making a choice from the menu).
3. Navigation—the participant uses the joystick to move from one area to the next, within the environment.

Before entering the café, participants were required to respond to three questions[5], all of which were presented verbally and textually by the computer. Participants were then able to complete the set list of tasks. The questions and a detailed breakdown of the tasks are included in the appendix.

Participants received minimal prompting throughout their use of the café environment. If

[5] One of these questions asked whether the user was in a wheelchair. This was not relevant to the present study, but was included in the environment because the Virtual Café had previously been used with students with learning and physical disabilities (Brown *et al.*, 1999).

participants asked about the next task, they were directed to read the checklist provided. Within tasks, prompts provided were of an indirect nature such that participants were not told exactly how to complete a particular task. For example, saying 'you need to eat your meal', rather than 'you need to click on the plate of food'.

Understanding Representations

Part A: Virtual Café. Just after participants had answered the three preliminary questions at the start of the café, and before they completed the set tasks on the list, participants were told that they were going to see some items in the café and would be asked to name them. The experimenter navigated around the café and pointed to the telephone, games machine, table, chairs, man and woman, in that order for all participants. After labelling the man and the woman at the bar, participants were asked 'What are they doing?' If participants responded by talking about individuals (e.g., he/she is talking), they were asked 'Who is he/she talking to?' Participants were then asked 'Are they doing anything else?' to elicit further ideas, until they had exhausted all suggestions.

Part B: Bar Video. Once participants had completed the set tasks in the café, they were shown a short clip of a video that showed a man and a woman standing at a bar. This was filmed in an empty bar, in order to match the conditions represented in the virtual café as closely as possible. Again the couple were standing facing each other, but were not actually talking. Participants were asked to label the same items as in the virtual café, in the following order: man, woman, telephone, table, chairs, games machine. The same questions were asked as before: 'What are they doing?' and 'Are they doing anything else?' Questioning took place approximately 30 minutes after participants had answered the same questions about the virtual café.

The video was always presented after participants had completed the questions and tasks in the café. This was to make sure that responses about the virtual café were not influenced by having previously answered questions about the video. That is, it may have been that answering questions about the video first prompted participants to answer in a similar way about the virtual scene, or cue them into a particular way of thinking. After completing this part of the procedure, pupils were thanked for their help (and given a sticker for their participation, where appropriate).

RESULTS AND DISCUSSION

Tests of Executive Function

Each of the five BADS sub-tests was scored according to procedures described in Wilson *et al.* (1986). Raw scores were then converted to a standardised profile score from 0 to 4, based on the performance of Wilson *et al.*'s control participants. Results for mean profile scores of each individual task (from a total of 4) are summarised in Fig. 2. For each task, the PIQ-match group scored highest and (except for the card shift test) the ASD group scored lowest. Pair-wise comparisons were used to

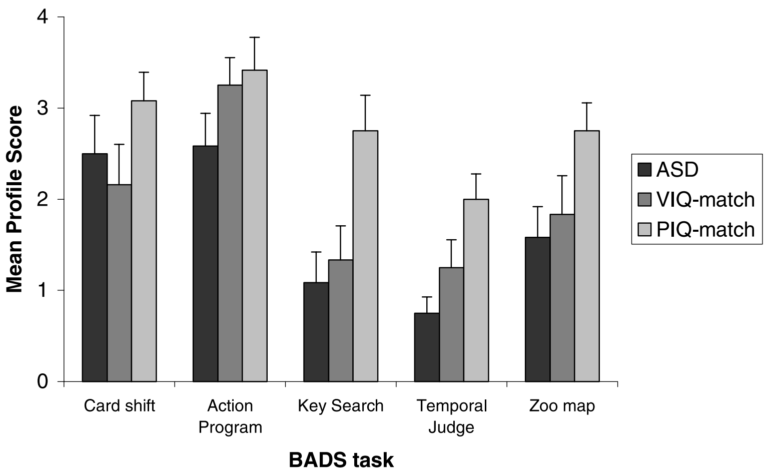

Fig. 2. Mean profile scores on BADS tasks according to group.

Table II. Mean Time in Seconds (and Standard Deviations) for Each Group to Complete Each Stage of the Training Environments, and Overall Mean Time for Completing the Combined Tasks in the Virtual Café

	Task				
Group	Training Trial 1	Training Trial 2	Training Trial 3	Training Trial 4	Total time in café*
ASD					206.1 (100.8)
Open space	58.7 (8.9)	55.6 (8.6)	56.9 (12.8)	49.8 (5.1)	–
Confined space	37.9 (17.4)	35.2 (10.9)	40.0 (24.2)	31.0 (7.8)	–
Search task	87.1 (38.7)	86.7 (52.7)	69.1 (25.0)	61.6 (26.4)	–
VIQ-match					262.36 (143.6)
Open space	57.3 (11.9)	52.6 (10.2)	51.6 (9.9)	54.0 (11.9)	–
Confined space	39.7 (17.5)	36.2 (13.9)	35.2 (8.7)	32.3 (5.3)	–
Search task	119.2 (92.6)	75.0 (19.0)	97.6 (56.7)	76.7 (31.7)	–
PIQ-match					169.8 (49.7)
Open space	54.0 (13.2)	49.8 (6.1)	50.2 (6.5)	48.4 (4.8)	–
Confined space	33.8 (14.3)	29.2 (5.2)	27.6 (3.7)	29.2 (5.5)	–
Search task	68.7 (34.8)	67.2 (28.7)	65.1 (21.0)	61.6 (19.0)	–

*Combined across all café tasks to give an overview of time taken.

determine whether there were differences between matched groups on each task. Given that the data for each task from the ASD group appeared in two separate comparisons, the significance level was adjusted to 0.025, using the Bonferroni correction.

There were no significant differences between the ASD and VIQ-match group on any of the individual tasks. In addition, the ASD group did not differ significantly from the PIQ-match group on the card shift test. However, the ASD group did differ significantly from the PIQ-match group, on three of the four remaining tasks: Action Program $t(11) = 3.46$, $p < .025$; Key Search $t(11) = 2.86$, $p < .025$; Temporal Judgment $t(11) = 3.56$, $p < .025$. On the final task, Zoo Map, a difference between the two groups approached significance: $t(11) = 2.54$, $p = .027$.

Despite a general trend across tasks for the PIQ-match group to gain higher profile scores than the ASD and VIQ-match groups, the shape of the profile for each group appears remarkably similar. Perhaps most strikingly, the tasks did not seem to pose a greater challenge for the ASD group than the VIQ-match group.

The 'Training' Virtual Environments[6]

The use of the training environments was divided into different stages: walking around the outside of the building (open space); walking through the corridors (confined space); and completing the task in the 'object room' (search task). Each stage was timed using a stopwatch. Participants were also classified according to whether they completed the search tasks correctly. This required selecting the correct objects (as shown on the stimulus card) and matching them in the correct order. Table II summarises the means and standard deviations for all groups, across the four trials in the training environment and the mean of the total time spent in the Virtual Café.

Open Space

Generally, the PIQ-match group were quicker than the other two groups at navigating around the building across the four trials. However, analysis of the data using a mixed ANOVA of 3(Group) × 4 (Trial), the second factor being a repeated measure, revealed that there was no significant difference between the groups and no significant interaction between Group and Trial. There was a significant main effect for Trial: $F(3,24) = 5.85$, $p < .01$. The mean times taken, across groups, decreased from Trials 1 to 4 suggesting improvement over time. This trend was confirmed using Page's trend test for both the ASD and PIQ-match groups, with $p < .01$ and $p < .05$, respectively. There was no significant trend of improvement in the VIQ-match group. The magnitude of the time reduction from Trials 1 to 4 was also significant, but only for the ASD group; $t(9) = 4.16$, $p < .005$. Correlations between mean time taken

[6] The value of n differs between some of the following statistical comparisons, due to some participants being reluctant to begin the training task by navigating around the outside of the building.

and performance on the BADS tasks, using the more conservative alpha level of 0.01, revealed only one significant result between mean time and Action Program profile score, in the VIQ-match group ($r = -.73$, $p = .01$). This indicates that individuals who took longer to navigate the outside of the building tended to gain a lower score on the Action Program task, which requires a physical manipulation of objects to reach a particular goal. It could be that this aspect of the task is related to the ability to manipulate the joystick, such that the environment is navigated successfully. Finally, correlations between IQ and time taken in this stage of the task, again with an alpha level of 0.01, showed no significant associations in any of the groups.

Overall, the ASD group improved to the same level as the PIQ-match group within four trials, whereas the VIQ-match group did not seem to benefit from practice to the same extent. The lack of any significant associations between time taken to complete the task, and background characteristics, suggest that good performance in the ASD group was not related with age, IQ or executive function abilities.

Confined Space

Participants in all groups tended to complete this section of the environment in very similar times. Again, using a mixed 3(Group) × 4 (Trial) ANOVA, there were no significant differences between any of the groups. There was also no significant main effect of Trial, and no Group × Trial interaction. Page's trend test showed that there were no significant trends of improvement across the four trials, within any of the groups and paired *t*-tests showed no significant differences between the first and last trials, for any of the groups. Correlations between background characteristics and time taken revealed only two significant associations. At an alpha level of 0.01, the first association was with chronological age of the VIQ-match group ($r = -.73$, $p < 0.01$), which suggests that older participants tended to complete this stage of the task in a shorter period of time, perhaps because older participants have had more experience with computers than younger members of the group. The second significant association was with Action Program on the BADS within the PIQ-match group ($r = -.81$, $p < .01$), which indicates that participants gaining a higher profile score on the Action Program task, tended to navigate the corridors more quickly. There were no significant associations in any of the groups between time taken and IQ.

There are a couple of reasons why participants may not have improved across trials. First, the program is set at a 'maximum speed', making it impossible to improve with practice if performance is already at, or near, ceiling. Alternatively, it may be that confined space in a VE is more difficult to navigate than open space. This is because the field of view is reduced, which tends to result in users being less able to consider their current orientation and how that relates to where they want to go (Neale, 2001).

Search Task

Generally, the PIQ-match group appeared to be at ceiling on this task, completing the trials quickly, and in similar times, on each of the four trials. The VIQ-match group showed an uneven profile, while the ASD group seemed to improve across the four trials. However, a mixed 3(Group) × 4 (Trial) ANOVA revealed that the groups did not differ significantly from each other in terms of overall time taken to complete the task. There was a main effect for Trial: $F(3,33) = 3.92$, $p < .05$, which was in the direction of progressive improvement (see Table II). There was no significant interaction between Group and Trial. There was a significant trend of improvement across trials for the ASD group, according to Page's trend test, at $p < .01$. There was no significant trend across trials for the VIQ- and PIQ-match groups. In addition, the magnitude of change from Trials 1 to 4, within the ASD group, was significant using a paired sample *t*-test: $t(11) = 3.11$, $p < .05$. This difference was not significant in either of the other two groups.

The correlations between executive function scores and background characteristics, and time taken to complete the search task, revealed no significant associations in the ASD group, or the VIQ-match group (at alpha level 0.01). There were a number of significant associations between background characteristics and mean time taken to complete the search task in the PIQ-match group. There was a significant relationship with VIQ ($r = -.71$, $p < .01$); PIQ ($r = -.75$, $p < .01$); FSIQ ($r = -.74$, $p < .01$) and Temporal Judgment ($r = -.71$, $p < .01$). In each case, the negative correlation suggests that individuals with higher scores on the standardised IQ and executive function tasks, took less time to complete the search task.

In this section of the training task, we were also interested in whether participants completed the search task correctly. All participants were at ceiling on matching the required objects in the specified order, without error. There was no occasion, within any of the groups, in which an incorrect object was chosen, or a correct object was moved to an incorrect position in the display bar.

Generally, the ASD group seemed to derive consistent, and significant, benefit from practicing the task over four trials. Noticeably, the VIQ-match group did not seem to improve across the trials, instead showing a more erratic pattern of performance. Significant associations in the PIQ-match group suggest that IQ and executive function were related with performance on the search task. Crucially, there were no associations between time taken and background characteristics in the ASD group. This suggests that the reduction in time for the ASD group, across the four trials, was not restricted to those with stronger executive function.

The Virtual Café

Time Taken to Complete Tasks

The time taken to complete all the tasks in the café was recorded (Table II). A one-way ANOVA was unable to identify a difference between groups. Correlations between background characteristics and time taken to complete the tasks in the café revealed a significant association with the Zoo Map task, in the ASD group ($r = -.78$, $p < .01$). This suggests that individuals in the ASD group with higher scores on the Zoo Map task, tended to complete the tasks in the café in a shorter period of time. One aspect of the Zoo Map task involves completing tasks in a clearly specified order, in much the same way as tasks are completed in a set order in the Café. This similarity could be reflected in the significant association between the two tasks, in the ASD group. There were no other significant associations between variables in either of the other two groups.

Completing the Tasks Successfully

The time taken to complete all the tasks in the café provides a simplified view of how participants completed this part of the procedure. A number of sub-tasks in the café could have proved difficult for participants, which would not be clear from the overall time taken (see Appendix for details of questions and tasks). However, for each of the main tasks in the café, the majority of participants—across all three groups—completed them without error and answered the questions correctly. There were a few exceptions to this, for example three members of the ASD group, three VIQ-match and two PIQ-match participants, inappropriately selected lager from the drinks menu, but all subsequently then chose a different drink when asked the next time. In addition, one participant in the ASD group responded 'yes' to the question 'Do you want a drink at the bar?' and proceeded to complete that stage of the café before the sequence at the black table.

The main difference between the groups when completing the tasks in the café was during the paying sequences (when participants could click on money on the screen to pay their bill, either with exact money or by paying too much and receiving change). Participants completed the paying sequence twice—once while at the table and once after ordering a drink at the counter. More members of the PIQ- and VIQ-match groups tended to supply the exact money the second time, perhaps because they had already received practice at the paying sequence. The participants in the ASD group did not seem to benefit from this practice, however, with the same number of participants requiring change the second time as at the paying sequence at the table. Overall, though, while a few members of the ASD group seemed to struggle with certain aspects of the café tasks, they did not seem especially impaired relative to the VIQ-match group, and the majority completed the tasks without error or the need for prompts.

Understanding Representations

All participants correctly identified the telephone, games machine, table, chairs, man and woman in the Virtual Café, as well as in the Bar Video. In response to the questions 'What are they doing?' and 'Are they doing anything else?', the main reply of interest (to either question) was whether participants said that the people were talking. Table III summarises the number of respondents in each group stating that the people were talking to the barman or to each other (combined across the two questions). Responses not included were statements like, 'just standing there', 'looking', 'moving their bodies', 'holding a drink', and 'waiting for something.' The majority of participants in each group said that the people at the bar were talking to

Table III. Number (and Percentage) of Participants in Each Group Saying that the People in the Virtual and Video Scenes were Talking

Group	Virtual Café	Bar Video
ASD (n = 12)	8 (67%)	9 (75%)
VIQ-match (n = 12)	10 (83%)	12 (100%)
PIQ-match (n = 12)	10 (83%)	11 (92%)

each other, or to the barman. Members of the ASD group were slightly less likely than the other groups to give this response, but not significantly so (based on χ^2 analyses).

In addition to investigating whether participants noticed the similarities between the two forms of representation, we also checked whether they acknowledged essential differences; being able to comment on sensible differences between the two forms of representation would mean that they did not consider them to be exactly alike. Therefore, on a follow-up occasion, we informally questioned our autistic participants about the differences between virtual reality and video. Ten of the twelve original group members were asked 'In what ways are video and virtual reality different from each other?' Some responses were very sophisticated, for example "Virtual reality is fictional and video is real"; " You can look around on virtual reality—it's more interactive. With a video you just have to sit and watch what happens". Others struggled to answer the question, which lead the experimenter to ask more focused questions, such as "Are there real people in virtual environments?" Seven out of the ten participants were able to acknowledge differences between virtual reality and video, even if only at the basic level of knowing that the people in virtual reality were not real. Of the remaining three participants, two thought that there were real people in the virtual scene, and one thought it was possible to move through a video with a joystick. Eleven out of the 12 VIQ-match participants were also questioned about differences between virtual reality and video. All stated that the people in VEs were not real, but that people in videos were. Overall, it seemed that the majority of participants questioned understood something about the representational nature of virtual environments.

Social Appropriateness

One of the interesting behaviors observed in pilot sessions with a group of adults with Asperger Syndrome, was the tendency to navigate between or very near to the couple standing at the bar in the Virtual Café, en route to buy a drink. This behavior seems socially inappropriate, especially as there was a large empty space at the bar, to the right of the couple (see Fig. 1(d)). We were keen to investigate whether this behavior would be shown by any of our participants in the main study and, if so, whether this effect would be more likely in the ASD group in particular (as suggested in an earlier pilot study).

The video segments of participants moving from the black table over to the bar were edited together on one tape, in a randomly generated order. This segment of video was not available for one VIQ-match participant, resulting in his matches from the other two groups being excluded from the following analyses also. Ten naïve raters (five males, five females) were asked to code each of the 33 video segments according to two criteria. The first was a categorical measure, which asked raters to code what they saw according to whether the user walked around, near, into (bumped) or between the couple at the bar. The second was a subjective measure, which asked raters to judge on a scale of 0–3 the extent to which the user intended to avoid the couple at the bar, with 0 being 'no intention to avoid' and 3 being 'clear intention to avoid'. All raters watched the video segments in the same order, either individually or in pairs. There was no discussion or collaboration allowed between the raters.

For the sake of brevity, and to avoid repetition of findings in relation to social appropriateness, only the results from the 'intention to avoid' ratings will be included below. The pattern of data from the categorical ratings, between matched pairs and for within group correlations, was almost identical. The only exceptions are where data from the categorical measure were used in additional analyses, and these are reported separately.

'Intention to Avoid' Ratings

The mean scores (from 0 to 3) across raters for this dimension are shown in Figs. 3 and 4, according to paired match in ascending VIQ and PIQ, respectively. A rating of 0 represents 'no intention to avoid' and a score of 3 represents 'clear intention to avoid'. In most cases, the variance is narrow and the error bars do not overlap, suggesting good agreement between raters, and a difference between the ASD participants and their matches.

Members of the ASD group with lower VIQs tended to be rated as having a lower intention to avoid compared both with their VIQ matches and the higher VIQ members of the ASD group. The data were submitted to a 2 (Group) × 11(Pair) mixed ANOVA, which showed a significant main effect for Group ($F(1,9) = 7.75$, $p < .05$) and Pair ($F(10,90) = 9.02$, $p < .001$), and a significant Group × Pair interaction ($F(10,90) = 19.9$, $p < .001$). The main effect for Group is in the expected direction with the ASD group rated as having less intention to avoid than the VIQ-match group.

A series of paired *t*-tests explored the nature of the significant interaction. The participants in pairs 1, 2, 4, 6 and 7 differed significantly from each other in the direction of the main effect. In each pair, the participant with an ASD was significantly more likely to be rated as having a lower intention to avoid the couple at the bar (Pair 1: $t(9) = 7.6$, $p < .001$; Pair 2: $t(9) = 5.2$, $p < .005$; Pair 4: $t(9) = 2.7$, $p < .05$; Pair 6: $t(9) = 4.8$, $p < .005$; Pair 7: $t(9) = 7.2$, $p < .001$). By contrast, two of the pairs at the upper end of the VIQ range differed significantly from each other in the opposite direction. The VIQ-matched participants in Pairs 8 and 10 were significantly more likely than their ASD counterparts to be judged as having a lower intention to avoid the couple at the bar (Pair 8: $t(9) = 7.6$, $p < .001$; Pair 10: $t(9) = 4.6$, $p < .005$).

The same 2(Group) × 11 (Pair) mixed measures analysis was applied to the data from the ASD group and their PIQ matches. There was a significant main effect for Group ($F(1,9) = 6.62$, $p < .05$) and Pair ($F(10,90) = 21.5$, $p < .001$). The effect for group was in the expected direction with a lower

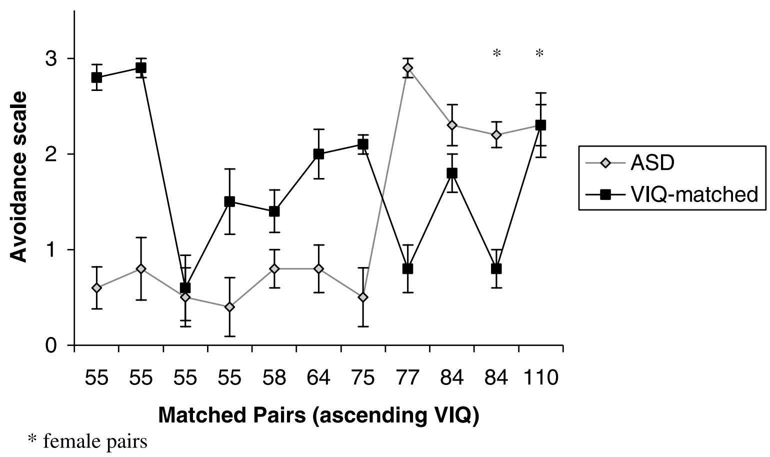

Fig. 3. Mean rating of 'intention to avoid' for pairs matched according to VIQ.

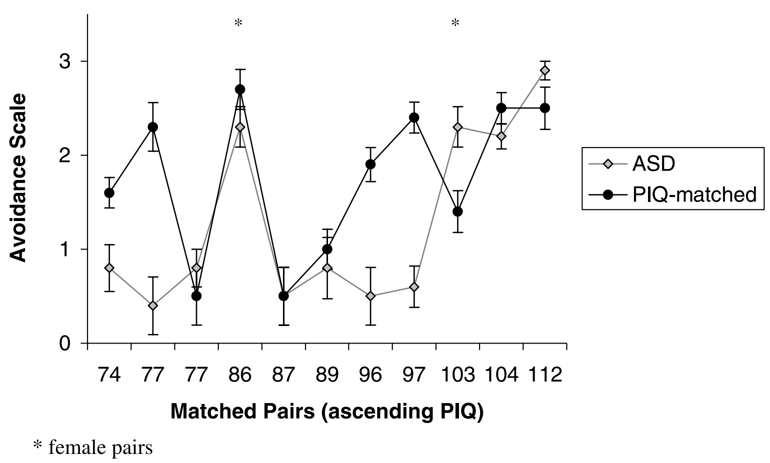

Fig. 4. Mean rating of 'intention to avoid' for pairs matched according to PIQ.

mean score in the ASD group compared to the PIQ group. There was also a significant Group × Pair interaction ($F(10,90) = 9.8$, $p < .001$). Paired comparisons revealed a similar, though less consistent, pattern to that found for the ASD and VIQ-matched groups. The ASD participants in Pairs 1, 2 and 7 were significantly more likely to be judged as having a low intention to avoid the people at the bar, compared to their PIQ-matches (Pair 1: $t(9) = 3.2$, $p < .05$; Pair 2: $t(9) = 3.6$, $p < .01$; Pair 7: $t(9) = 4.1$, $p < .005$). By contrast, the ASD participants in Pairs 8 and 9 were significantly less likely to be judged as having a low intention to avoid, compared to their PIQ-matches (Pair 8: $t(9) = 6.2$, $p < .001$; Pair 9: $t(9) = 3.2$, $p < .05$).

There were no significant correlations between scores on individual BADS tasks and 'intention to avoid' ratings for any of the groups, but there were significant associations between 'intention to avoid' and IQ, but only in the ASD group: VIQ ($r = .76$, $p < .01$; PIQ ($r = 0.61$, $p < .05$; FSIQ ($r = 0.85$, $p < .001$; $df = 9$ in all cases).

Additional Analyses Using the Categorical Ratings

We also checked for an association between participants' general tendency to bump into things in the virtual environment and the raters' judgements of whether they bumped into the couple. We took a measure of 'movement control' that consisted of the number of times each participant bumped into the wall of the corridor on the fourth training trial. This trial directly preceded the café environment, thereby providing the most relevant measure of participants' ability to control their movement through the environment. The mean number of bumps into the corridor for each group were as follows: ASD = 1.1 (SD = 0.94); VIQ-match = 2.2 (SD = 1.8); PIQ-match = 1.2 (SD = 1.2). Correlations between number of bumps in the corridor and number of raters judging that the participant bumped into the people at the bar, revealed no significant associations within any of the three groups. In other words, a general measure of movement control did not predict a judged tendency to bump into people at the bar.

GENERAL DISCUSSION

This study aimed to systematically investigate the use and understanding of virtual environments by people with ASDs in the 'non retarded' full-scale IQ range (DSM-IV, 1994), compared to closely matched control groups. Participants practised using the equipment over a series of four training trials, and also completed a number of tasks in a Virtual Café. Time spent completing tasks was examined, as well as any errors made. We also included some preliminary investigations into understanding VEs as having representational qualities and considered the social appropriateness of participants' performance.

Given the lack of information available in the background literature regarding the performance of people with ASDs in virtual environments, we were not sure what to expect from the present group. Overwhelmingly, however, the picture of their ability to use the technology is positive. Participants with ASDs learned to use the equipment quickly and showed significant improvements in performance after only a few trials in the training environment. The VIQ-matched group was different in this respect, tending to show a less consistent pattern of performance and failing to demonstrate significant time improvements across the four trials. The performance of the ASD group is even more striking in the context of their weak executive function abilities, as assessed on standardised tasks. In order to successfully complete tasks in the virtual environments, participants needed to follow instructions, inhibit incorrect responses, keep track of goals and co-ordinate use of equipment to achieve specified aims. While there were links between speed of performance and executive function in the VIQ-match group, the same was not true for those with ASDs. Despite weaker executive function abilities than the VIQ-match group (although not significantly so), the participants with ASDs outperformed the VIQ-match group in terms of making improvements across trials on time taken to complete tasks.

It is possible that the structuring of tasks within the environments provides extra support for people with weaker executive abilities, thereby supporting performance on the tasks in the virtual environments and preventing any association with the standardised BADS tasks. However, it is not clear why this effect should operate within the ASD group only. There were a number of significant associations between executive function scores and time taken to complete different stages of the training environment within the VIQ- and PIQ-match groups, suggesting that there were shared demands between some of the BADS tasks and performance in the training environment. Nevertheless, it remains likely that the supportive structure of the environments, including the checklist of tasks, enabled participants to focus on the main tasks and complete them correctly. Future virtual environments should be built with this in mind, in order to provide supportive situations in which participants can concentrate on target skills, rather than struggling to overcome problematic executive demands. Additionally, gradually reducing the level of support and structure provided over subsequent sessions could help to ensure that built-in structure and prompts do not lead to 'unrealistic' scenarios in which responses could be rote-learned. The presentation of scenes, and questions/tasks within them could also differ slightly each time they are presented to increase the flexibility of scenarios and encourage participants to think through the situation each time (Parsons & Mitchell, 2002).

The majority of participants with ASDs experienced few difficulties completing tasks in a Virtual Café. Negotiating the main tasks in the café entailed a relatively complex series of sub-tasks, such as moving to the correct region of the room, answering questions correctly and clicking on food in order to 'eat' before proceeding to the next stage. While a few individuals in the ASD and VIQ-matched groups required support from the experimenter to complete some of the tasks (e.g., the paying sequence), this was not the case for the majority. In most cases, the participants with ASDs were not different from the PIQ-matched group in terms of ability to complete the tasks in the café in the specified order, without help from the experimenter. Overall, the use of the virtual environments by people with ASDs was on a par with a group of normally developing PIQ-matched individuals.

A second main strand of investigation was whether participants understood the virtual environments as having basic representational qualities. Encouragingly, the majority of participants with ASDs in the present study interpreted the virtual environment in a nonliteral way by suggesting that the virtual figures standing at the bar were talking to each other when, in fact, they were standing facing each other in close proximity, but not talking. This suggests that participants imbued the figures with 'people-like' interpersonal behavior. In addition, participants responded in a similar way to video footage of some real people in a real bar, suggesting that the two scenes were treated analogously. This is reinforced by anecdotal evidence from some members of the ASD group. For example, one female said "It looks just like a real café!" when she first saw the virtual café, and one of the males said "I just sat on those seats didn't I?", when shown the video of the bar, after completing the tasks in the virtual café. This seemingly good understanding of how VEs relate to the real world is especially encouraging given that some of the actions, and interactions, in the virtual café were not particularly realistic. For example, participants could choose items and 'speak' to characters in the environment by clicking on limited choices depicted in text boxes, which is quite different from the way similar tasks would be tackled in the real world.

Additionally, participants were able to verbalize basic differences between virtual environments and videoed scenes. The majority of participants in the ASD group stated that the people in the virtual café were not real, but that people in the video were real. This suggests that despite answering questions about the video and the virtual environment in a similar way, participants were not thinking of them as exactly the same. It seems that the majority of the present group of participants were able to make a distinction between reality and representation, at least at this very basic level. Further evidence for the ability of autistic participants to distinguish representation from reality has been reported in Parsons and Mitchell (1999). Children with autism were able to make correct judgments about the content of a story protagonist's belief, with the aid of pictorial thought bubbles, whilst also judging correctly about the current state of reality. Nevertheless, it is worth noting that two members of the ASD group in the present study said there were real people in the virtual café, suggesting a need for caution when using the technology with some members of this clinical group.

The final area of investigation followed up an observation from a pilot group of participants with ASDs, who showed a tendency to navigate very near to the people at the bar when ordering a drink, rather than moving to the large, clear space to the right of the couple. Ratings from naïve coders who watched video clips of participants' navigation from the table to the bar, revealed an interesting pattern of results when performance was compared between members of matched pairs. Participants in the ASD group, especially those with lower VIQs (from 55 to 75), were more frequently rated as walking between, or bumping into, the people at the bar, and more likely to be judged as having a 'low intention to avoid' the couple, compared to their VIQ-match counterparts. Interestingly, the reverse pattern was seen in some of the pairs with higher VIQs. Similar results were found when ASD participants were compared to their PIQ-matched counterparts, although the pattern was not as consistent as with the VIQ-match group. Significant correlations in the ASD group between these ratings and IQ measures support these findings.

There are a number of possible explanations for this observed tendency. First, there could be a general motor difficulty in executing the movements necessary to avoid the couple. There is some evidence that people with ASDs are impaired on tests of fine and gross motor control, relative to normal controls (Ghaziuddin & Butler, 1998; Ghaziuddin, Butler, Tsai, & Ghaziuddin, 1994). It might be that participants know they should avoid the couple at the bar, but are unable to do so. However, our general measure of movement control, which consisted of the number of bumps each participant made into the corridor wall on the fourth training trial, showed no significant correlation with the number of raters who judged the participant as bumping into the couple. Indeed, the ASD group had the lowest mean number of bumps into the corridor walls, compared to the other two groups. It seems that a motor difficulty in executing the correct movement is an unlikely explanation for the observed findings. However, it remains a possibility that participants may have a particular difficulty navigating around freestanding obstacles, like people or objects, and not with moving past 'fixed' spaces like walls. Future research could include direct comparisons of navigation around different objects, using different kinds of navigational devices (e.g., keyboard direction keys, game keypad), to provide a more complete test of navigational ability.

Second, it could be that participants in the ASD group bumped into the people at the bar because they did not interpret the figures as representations of people. This explanation also seems unlikely given that the majority of participants seemed to have some basic insight into the representational quality of the virtual environment, and seemed to interpret the figures as having 'people-like' qualities. It is possible, though, for participants to understand the representational nature of the figures, yet still relate to people as if they were inanimate objects. This could signal a failure to understand the special status of people in the social world (cf. Hobson, 1993). A third explanation might be that executive weaknesses in planning the movement necessary to avoid the couple at the bar contributed to the observed effect. Again, this seems unlikely due to the fact that there were no significant associations between scores on tests of executive function and participants' judged tendency to bump into the people at the bar.

A remaining possibility is that some participants with ASDs (especially those in the lower VIQ range) have a weak understanding of appropriate behavior concerning the personal space of people in social situations, at least as presented in a virtual environment. It could be argued that the notion of personal space is not seen (by people with ASDs, or anyone else) as applying to representations of people in a virtual scene. However, the finding that the majority of our non-autistic participants made an effort to avoid the couple at the bar, suggests otherwise. Anecdotally, teachers of the participants included in the present study reported poor understanding of personal space, and inappropriate social contact, amongst members of the group. This may apply only to those with lower VIQs, since participants with higher VIQs tended to exhibit few problems in avoiding the people at the bar. On the other hand, it might be that the difficulty with personal space is a relatively coarse measure of social understanding, which is not sensitive enough to elicit errors from the higher-functioning participants.

Evidence is mixed regarding the extent to which IQ is linked to social behavior and understanding. People with lower verbal abilities are less likely to pass tests of false belief (Happe, 1995; Sparrevohn & Howie, 1995), and links have been found between an individual's ability to pass false belief tasks and their competence in everyday behaviors (Fombonne, Siddons, Achard, Frith, & Happe 1994; Frith, Happe, & Siddons, 1994). However, other studies have found no association between IQ and developments in social understanding in groups of children and adolescents with autism (Berger *et al.*, 1993; Freeman, Del'Homme, Guthrie, & Zhang, 1999). The finding that participants with ASDs in the present study, with lower IQs, were more likely to err on a task involving a social judgement suggests there may be a link between IQ and social understanding. Of course, this was in relation to a virtual, rather than real-world, context; thus, it remains an open question whether IQ has a significant association with social understanding.

Nevertheless, some people with ASDs might have difficulty understanding the norms and expectations governing the acceptable negotiation of personal space. This lack of understanding may be specific to social situations as depicted in virtual reality, or it might be indicative of a deeper difficulty with social spaces in real life. If the latter, virtual reality may offer the ideal opportunity for participants to learn about the appropriateness of certain aspects of their behavior, in a safe and supportive environment. Clearly, there is much potential for future research to investigate the extent and nature of this difficulty. In particular, investigations into whether participants' sensitivity to personal space within the VE is linked to real-world behavior would prove extremely insightful, as would the possibility of whether the present results could be replicated with a larger sample of participants.

Finally, although the results for the majority of participants with ASDs were positive, a few seemed to struggle with the tasks, and appeared to have difficulty understanding the virtual environment as a representation of reality. As with most interventions/packages/styles of learning—especially within this very heterogeneous clinical population—this technology may not be suitable for everyone. Encouragingly, though, most of the participants with ASDs in the present study were as proficient as their PIQ-matched counterparts in terms of their use of the virtual environments. They also seemed to understand the virtual café as appropriately representational. This provides a promising starting point from which the future development of virtual technology for people with ASDs can build upon and improve.

ACKNOWLEDGMENTS

This research is funded by the Shirley Foundation. Many thanks to the pupils and staff from Rosehill, Foxwood, Brackenfield and The Bluecoat, Schools, all in Nottingham, U.K. Thanks also to the National Autistic Society, U.K., the rest of the 'AS Interactive' research team, and to David Moore, Thusha Rajendran and Cheryl Trepagnier for their comments on an earlier draft. More information about the project can be found at www.asinteractive.org.uk.

Appendix. Summary of BADS Tasks

Task	Description
Rule Shift Cards Test	Measures the ability to shift from one rule for responding to another; requires participants to keep track of the previous stimulus and the present rule.
Action Program Test	Requires participants to implement a plan of action in order to solve a novel, practical problem. Involves the physical manipulation of apparatus, rather than using a pen and paper to achieve a solution.
Key Search Test	A pen and paper task which requires the participant to plan an efficient and effective course of action. Participants can monitor their own performance by seeing how the line they draw on the paper helps them to achieve the overall aim of the task.
Temporal Judgment Test	Asks participants to make a 'best guess' to four questions concerned with the period in time of certain events, e.g., How long do most dogs live for? How long is a routine dental check-up?
Zoo Map Test	A pen and paper task with 'low demand' and 'high demand' versions. The 'high demand' version requires participants to plan a route around a zoo in order to visit a number of specified places, whilst observing certain rules. This assesses planning ability, as well as ability to modify responses based on feedback. The 'low demand' version requires participants to follow a set of instructions in order to visit the same places.

Breakdown of Tasks in the Virtual Café

	Navigations	Communication	Interaction
A. Enter details		***Are you male or female?***	Click on answers; **Male** or **Female**
		Are you over 18?	Click on answers; **Yes** or **No**
		Are you in a wheelchair?	Click on answers; **Yes** or **No**
B. Find a table	Move to table	*Would you like to sit here?*	Click on answers; **Yes** or **No**
C. Order food and drinks			Click on menu; Click on food Click on drink
Waiter brings food			Click no food to eat it
		Waiter walks over and says *Was everything OK? Here is the bill.*	
D. Pay for food and drink			Click on coins or notes; Click on green OK arrow
D1. If correct money is given		Waiter says *Well done. You gave me the right amount of money.*	
D2. If change is needed		Waiter says [how much change you have].	
D3. If not enough money is paid		Waiter says *sorry. Its not enough money. I need [x amount].*	Click on coins; Click on green OK arrow
		Waiter says *Well done. You gave me the right amount of money.*	
E. Order a drink from the bar	Move over to the bar	Barman says *Do you want a drink at the bar?*	Click on answers; **Yes** or **No**
		Barman says *Have you been served?*	
A menu of drinks appears			Click on choice of drink from menu
Drink appears on bar			Click on the drink to drink it.
		Barman says *Was everything OK? Here is the bill.*	
F. Pay for the drink (exactly same as section D)			

Note: Verbal prompts in italics, textual prompts in bold.

REFERENCES

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

Berger, H. J. C., van Spaendonck, K. P. M., Horstink, M. W. I. M., Buytenhuijs, E. L., Lammers, P. W. J. M., & Cools, A. R. (1993). Cognitive shifting as a predictor of progress in social understanding in high-functioning adolescents with autism: a prospective study. *Journal of Autism and Developmental Disorders, 23*, 341–359.

Bernard-Opitz, V., Sriram, N., & Nakhoda-Sapuan, S. (2001). Enhancing social problem solving in children with autism and normal children through computer-assisted instruction. *Journal of Autism and Developmental Disorders, 31*, 377–384.

Blade, R. A., & Padgett, M. L. (2001). Virtual environments standards and terminology. In K. Stanney (Ed.), *Handbook of Virtual Environments: Design, Implementation and Applications*. pp. 15–27. New Jersey: Lawrence Erlbaum Associates.

Brown, D. J., Neale, H. R., Cobb, S. V. G., & Reynolds, H. (1999). Development and evaluation of the virtual city. *International Journal of Virtual Reality, 4*, 28–41.

Charman, T., & Baron-Cohen, S. (1992). Understanding drawings and beliefs: A further test of the metarepresentation theory of autism. *Journal of Child Psychology and Psychiatry, 33*, 1105–1112.

Chen, S. H. A., & Bernard-Opitz, V. (1993). Comparison of personal and computer-assisted instruction for children with autism. *Mental Retardation, 31*, 368–376.

Clancy, H. (1996). Medical field prescribes virtual reality for rehabilitation therapy. *Computer Reseller News, 698*, 76.

Cobb, S., Nichols, S. C., Ramsey, A., & Wilson, J. (1999). Virtual reality induced symptoms and effects (VRISE). *Presence: teleoperators and virtual environments, 8*, 169–186.

Cromby, J. J., Standen, P. J., & Brown, D. J. (1996). The potentials of virtual environments in the education and training of people with learning disabilities. *Journal of Intellectual Disability Research, 40*, 489–501.

Eynon, A. (1997). Computer Interaction: An update on the AVATAR program, *Communication*, Summer, 1997, p. 18.

Fombonne, E., Siddons, F., Achard, S., Frith, U., & Happe, F. (1994). Adaptive behavior and theory of mind in autism. *European Child and Adolescent Psychiatry, 3*, 176–186.

Freeman, B. J., Del'Homme, M., Guthrie, D., & Zhang, F. (1999). Vineland adaptive behavior scale scores as a function of age and initial IQ in 210 autistic children. *Journal of Autism and Developmental Disorders, 29*, 379–384.

Frith, U., Happe, F. G. E., & Siddons, F. (1994). Autism and theory of mind in everyday life. *Social Development, 3*, 108–123.

Ghaziuddin, M., & Butler, E. (1998). Clumsiness in autism and Asperger syndrome: A further report. *Journal of Intellectual Disability Research, 42*, 43–48.

Ghaziuddin, M., Butler, E., Tsai, L., & Ghaziuddin, N. (1994). Is clumsiness a marker for Asperger syndrome? *Journal of Intellectual Disability Research, 38*, 519–527.

Happe, F. G. E. (1995). The role of age and verbal ability in the Theory of Mind task performance of subjects with autism. *Child Development, 66*, 843–855.

Heimann, M., Nelson, K., Tjus, T., & Gilberg, C. (1995). Increasing reading and communication skills in children with autism through an interactive multimedia computer program. *Journal of Autism and Developmental Disorders, 25*, 459–480.

Hobson, R.P. (1993). *Autism and the Development of Mind*. Hove: Lawrence Erlbaum Associates.

Hughes, C., & Russell, J. (1993). Autistic children's difficulty with mental disengagement from an object: its implications for theories of autism. *Developmental Psychology, 29*, 498–510.

Hughes, C., Russell, J., & Robbins, T. W. (1994). Evidence for executive dysfunction in autism. *Neuropsychologia, 32*, 477–492.

Leekam, S. R., & Perner, J. (1991). Does the autistic child have a metarepresentational deficit? *Cognition, 40*, 203–218.

Mitchell, P., Saltmarsh, R., & Russell, H. (1997). Overly literal interpretations of speech in autism: Understanding that messages arise from minds. *Journal of Child Psychology and Psychiatry, 38*, 685–691.

Moore, D. J. (1998). Computers and people with autism/asperger syndrome. *Communication*, 20–21.

Moore, M., & Calvert, S. (2000). Brief report: Vocabulary acquisition for children with autism: teacher or computer instruction. *Journal of Autism and Developmental Disorders, 30*, 359–362.

Moore, D. J., McGrath, P., & Thorpe, J. (2000). Computer aided learning for people with autism—a framework for research and development. *Innovations in Education and Training International, 37*, 218–228.

Moore, D., & Taylor, J. (2000). Interactive multimedia systems for students with autism. *Journal of Educational Media, 25*, 169–177.

Neale, H. (2001). Virtual environments in special needs education: Considering users in design. Unpublished PhD Thesis. University of Nottingham, U.K.

Nichols, S. C. (1999). Virtual reality induced symptoms and effects (VRISE): Methodological and theoretical issues. Unpublished PhD Thesis. University of Nottingham, U.K.

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

Ozonoff, S., Strayer, D. L., McMahon, W. M., & Filloux, F. (1994). Executive function abilities in autism and tourette syndrome: An information processing approach. *Journal of Child Psychology and Psychiatry, 35*, 1015–1032.

Parsons, S., Beardon, L., Neale, H. R., Reynard, G., Eastgate, R., Wilson, J. R., Cobb, S. V. G., Benford, S. D., Mitchell, P., & Hopkins, E. (2000). Development of social skills amongst adults with asperger's syndrome using virtual environments: the 'AS Interactive' project. In P. Sharkey, A. Cesarani, L. Pugnetti & A. Rizzo (Eds.) *3rd ICDVRAT, Sardinia Italy*. University of Reading, pp. 163–170.

Parsons, S., & Mitchell, P. (1999). What children with autism understand about thoughts and thought bubbles. *Autism, 3*, 17–38.

Parsons, S., & Mitchell, P. (2002). The potential of virtual reality in social skills training for people with autistic spectrum disorders. *Journal of Intellectual Disability Research, 46*, 430–443.

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

Rothbaum, B. O., & Hodges, L. F. (1999). The use of virtual reality exposure in the treatment of anxiety disorders. *Behavior Modification, 23*, 507–525.

Silver, M., & Oakes, P. (2001). Evaluation of a new computer intervention to teach people with autism or Asperger syndrome to recognize and predict emotions in others. *Autism, 5*, 299–316.

Sparrevohn, R., & Howie, P. M. (1995). Theory of mind in children with autistic disorder: Evidence of developmental progression and the role of verbal ability. *Journal of Child Psychology and Psychiatry, 36*, 249–263.

Strickland, D. (1996). A virtual reality application with autistic children. *Presence: Teleoperators and Virtual Environments, 5*, 319–329.

Strickland, D., Marcus, L. M., Mesibov, G. B., & Hogan, K. (1996). Brief report: Two case studies using virtual reality as a learning tool for autistic children. *Journal of Autism and Developmental Disorders, 26*, 651–659.

Swettenham, J. G. (1996). Can children with autism be taught to understand false belief using computers? *Journal of Child Psychology and Psychiatry, 37*, 157–165.

Trepagnier, C. G. (1999). Virtual environments for the investigation and rehabilitation of cognitive and perceptual impairments. *NeuroRehabilitation, 12*, 63–72.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. The Psychological Corporation.

Wilson, B. A., Alderman, N., Burgess, P. W., Emslie, H. C., & Evans, J. J. (1986). *Behavioral Assessment of the Dysexecutive Syndome*. Thames Valley Test Company: Flempton, Bury St. Edmunds.

Wilson, P. N., Foreman, N., & Stanton, D. (1998). A rejoinder. *Disability and Rehabilitation, 20*, 113–115.