Virtual environments and autism: a developmental psychopathological approach

Rajendran, Gnanathusharan

Published in:
Journal of Computer Assisted Learning

DOI:
10.1111/jcal.12006

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Link to publication in Heriot-Watt University Research Portal

Citation for published version (APA):
Virtual environments and autism: a developmental psychopathological approach

G. Rajendran
Heriot-Watt University, Edinburgh, UK

Abstract
Individuals with autism spectrum disorders supposedly have an affinity with information and communication technology (ICT), making it an ideally suited media for this population. Virtual environments (VEs) – both two-dimensional and immersive – represent a particular kind of ICT that might be of special benefit. Specifically, this paper discusses the importance of psychological theory for VE designed for this population. I describe the contribution that different theories of autism (e.g., theory of mind, executive function, weak central coherence theory) have made and can make, as well as the potential of other non-autism-specific theories (e.g., embodied cognition). These technologies not only illuminate our understanding of autism, but they can also be used to develop new technologies for people without autism. So, as well as being an area of specialism, I argue that VE research in autism has extended – and will go onto – the boundaries of human–computer interaction more generally. This is because autism provides a unique window into human social communication and learning. Further, this field offers a chance for better inclusivity for individuals with autism within a digital society.

Keywords
autism, autism spectrum disorders, ASD, virtual environments, developmental psychopathology

Introduction
Despite being around since the mid-1990s, the evidence base for virtual environments (VEs) for individuals with autism spectrum disorders (ASD) is relatively small. Hence, it remains a rich and challenging area, with still many unknowns; so there are good grounds for exploring ideas. I hope to illustrate this array of exciting avenues for research and application in this paper. I begin with a brief overview of autism and information and communication technology (ICT) because this area has become a research specialism in itself and also VE represents a particular kind of ICT that might be especially beneficial for people with autism. I then go on to discuss how different theories provide new ways of understanding and treating autism using VE. Finally, I conclude with more speculative thoughts about what the future might hold.

Throughout the paper, I take a developmental psychopathological viewpoint (Cicchetti, 1984). From this perspective, it is assumed that we can learn more about the typical functioning of an organism by the study of its pathology and, similarly, more about its pathology by studying its typical condition (Cicchetti, 1984). So, I argue that the relationship between people with autism and ICT is important for at least three reasons. First, it can tell us a great deal about the interests (and aversions) of people with this condition. Second, it offers a myriad of opportunities to intervene, to support, and facilitate skills development in people with autism. Third, research into ICT and autism benefits our understanding of human–computer interaction more generally.
I also use the term VE to include virtual reality (VR), which usually involves a level of three-dimensional (3D) representation. In recent review of the applied work in this area, Parsons and Cobb (2011) critique the studies to date of VR and autism. They conclude that despite limited research, VR does have unique potential for children with autism. Here, rather than reiterate their points, I will focus on the role of psychological theory in VEs for individuals with autism.

**Brief overview of autism**

ASD (henceforth autism) are currently characterized as a group of syndromes: autism *per se* (sometimes known as autistic disorder), Asperger syndrome, and pervasive developmental disorders—not otherwise specified – in Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition, Text Revision) [American Psychiatric Association (APA), 2000]. All three fall under the overarching category of pervasive developmental disorders (APA, 2000). Autism is currently diagnosed and conceptualized as a triad of impairments in socialization, communication and repetitive interests (APA, 2000). However, the nosology of autism seems likely to change with the advent of Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition), in which the autism spectrum looks set to become unified as an umbrella term (APA, 2000). Autism is currently diagnosed and conceptualized as a triad of impairments in socialization, communication and repetitive interests (APA, 2000). However, the nosology of autism seems likely to change with the advent of Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition), in which the autism spectrum looks set to become unified as an umbrella term (APA, 2000).

Notwithstanding any diagnostic changes, theoretical accounts have tried to explain the paradox of how this triad can be expressed in different individuals of differing ages and with different intellectual abilities, and within the same individual across their lifetime. Despite decades of research, however, there is no definitive answer to what is core to the disorder other than that it is developmental. Over the past 25 years or so, the cognitive theories of autism [theory of mind, executive dysfunction, weak central coherence (WCC)] have been hugely influential in understanding different aspects of autism (see Rajendran & Mitchell, 2007, for a review). However, the general consensus seems to be that a single unifying theory of autism may not be possible, but a multifactorial account may be the most parsimonious one (Happé, Ronald, & Plomin, 2006; Pellicano, Maybery, Durkin, & Maley, 2006; Pennington, 2006; Rajendran & Mitchell, 2007). Yet, even this may not be whole story given the lack of stability of performance over time in tasks tapping into these three cognitive processes (Pellicano, 2010).

**Brief overview of theories of autism**

Autism is perhaps, therefore, best conceptualized as a multifaceted condition: a disorder of social cognition, of cognition, of emotion, of perception and of movement. I will return to all these themes in greater detail under the auspices of the different theories. Importantly, these theories are not mutually exclusive and each has a role in illuminating autism in its totality (Rajendran & Mitchell, 2007). The theory of mind hypothesis of autism has perhaps been the most influential theory in recent years. In essence, this theory states that individuals with autism fail to ‘impute mental states to themselves and others’ (Premack & Woodruff, 1978, p. 515) and that this deficit shows itself as inability to mentalize, or failure to take into account others’ mental states. The theory of mind hypothesis can therefore explain some of the most profound repercussions of autism: from difficulties in making friends (often leading to loneliness) to problems understanding that what people say can be different from what they mean (e.g., nonliteral language such as sarcasm and figures of speech).

As a developmental precursor to theory of mind, the enactive mind hypothesis (Klin, Jones, Schultz, & Volkmar, 2003) posits that from infancy onwards the autistic mind is not attuned to social stimuli in the same way as the neurotypical mind; for example, following another’s eye gaze and a preference for biological over mechanical motion. Klin et al. (2003) argue these early developmental differences lead to downstream problems in the development of theory of mind and other aspects of social cognition. Specifically, joint attention is the mechanism by which eye gaze is used to follow or direct another person’s attention to a point of reference (Kim & Mundy, 2012); for example, following another’s eye gaze and a preference for biological over mechanical motion. Klin et al. (2003) argue these early developmental differences lead to downstream problems in the development of theory of mind and other aspects of social cognition. Specifically, joint attention is the mechanism by which eye gaze is used to follow or direct another person’s attention to a point of reference (Kim & Mundy, 2012); for example, following another’s eye gaze and a preference for biological over mechanical motion. Klin et al. (2003) argue these early developmental differences lead to downstream problems in the development of theory of mind and other aspects of social cognition. Specifically, joint attention is the mechanism by which eye gaze is used to follow or direct another person’s attention to a point of reference (Kim & Mundy, 2012); for example, following another’s eye gaze and a preference for biological over mechanical motion. Klin et al. (2003) argue these early developmental differences lead to downstream problems in the development of theory of mind and other aspects of social cognition. Specifically, joint attention is the mechanism by which eye gaze is used to follow or direct another person’s attention to a point of reference (Kim & Mundy, 2012); for example, following another’s eye gaze and a preference for biological over mechanical motion.
how children take another’s knowledge state into consideration in being an effective communicator.

In contrast to the above theories, the executive function (EF, e.g., Russell, 1996) theory of autism did not arise from neurotypical research; rather, its conception came from researchers who noted that some symptoms of autism were similar to those associated with specific brain injury in adults (particularly to the frontal lobes). EF is an umbrella term for behaviours needed for goal-directed behaviour (e.g., planning, cognitive flexibility, inhibition, etc.), so EF theory is particularly useful when trying to understand the autistic need for sameness, a difficulty switching attention and lack of impulse control. An additional strength of EF theory – which is only beginning to be realized – is that it can also explain some of the motor and movement self-regulatory aspects of autism (Rajendran & Mitchell, 2007).

A key strength of WCC theory (Frith, 1989) is that it explains some of the non-social (as well as the social) features of autism, such as the attention to acute detail that ranges from pedantry to obsession. The essence of the theory is that typically developing individuals process information by extracting the overall meaning or gist. Whereas, the ‘cognitive style’ of individuals with autism is the opposite: to process the constituent parts rather than the global whole. Another important aspect of this theory is that it is non-pejorative, in the sense that it makes predications about better performance in individuals with autism. Whereas the theory of mind hypothesis of autism, by contrast, is based on a deficit rather than difference conceptualization of autism – and, therefore, makes predications about worse performance in individuals with autism.

Although WCC is a discrete theory in itself, other theories have come in its wake to also try and explain perpetual processing in autism. For example, reduced generalization theory (Plaisted, 2001) states that individuals with autism have reduced processing of the similarities that are held between stimuli and situations – and this can explain why generalizing learning from one situation to another can be so hard for individuals with autism. Enhanced perceptual processing (Mottron, Dawson, Soulieres, Hubert, & Burack, 2006) makes an even stronger claim than WCC theory by suggesting that across-the-board perceptual processing is generally better in individuals with autism; for example, in individuals with autism with savant syndrome (Mottron & Belleville, 1993).

Finally, I will discuss embodied cognition: how psychological processes are influenced by the body (Glenberg, 2010). Recently, movement researchers have argued that understanding the motoric aspects of autism helps us also to understand the cognitive aspects of autism (Izawa et al., 2012). Here, I argue that as ICT devices become more movement based (e.g., Microsoft Kinect), this can be exploited to both better understand autism and help design the next generation of human–computer interfaces.

**Autism and ICT**

Within the last decade or so, ICT and autism have become a research specialism in itself arguably because ICT seems intrinsically less socially threatening (Swettenham, 1996; Trepagnier, 1999; Rajendran & Mitchell, 2000; Parsons & Mitchell, 2002; Rajendran, Mitchell, & Rickards, 2005; Goodwin, 2008). Additionally, in certain contexts the slower pace and asynchronous aspects of some ICT interfaces (e.g., email, text chat/instant messaging) have been argued to be more in tune with the autistic style of learning and interaction (e.g., Rajendran & Mitchell, 2006). This has led to arguments that ICT can provide not only a route into the social world for individuals with autism (Rajendran & Mitchell, 2006; Durkin, 2010; Jones, 2010), but also controversies.

‘Aren’t you just making people with autism more autistic using ICT?’

One oft-voiced criticism is that ICT is just going to make an already socially isolated person more withdrawn, more autistic and less likely to interact with people (Howlin, 1998; Latash, 1998). At its most extreme, Greenfield (Swain, 2011) has gone as far as saying the Internet might even cause autism. However, Bishop (2011) makes a strong case against this.

There are two arguments against this position. First, there is no evidence to suggest that ICT exacerbates the social-communicative problems in autism (or any other aspect of autism for that matter). Second, ICT often facilitates interactions with others either directly (because the person with autism tries to gain the attention of someone nearby, e.g., Alcorn et al., 2011, or via
the help of a teacher, e.g., Parsons, Leonard, & Mitchell, 2006), or indirectly because they are communicating with another person via ICT (either via a shared computer, e.g., Rajendran et al., 2005, or via a network, e.g., Rajendran & Mitchell, 2006; Parsons, Millen, Garib-Penna, & Cobb, 2011). Indeed, the importance of a facilitator, be it a parent or a teacher, cannot be understated in helping to structure ICT interactions. In short, working with ICT increases rather than reduces opportunities for communication between the child and the adult (Porayska-Pomsta et al., 2012).

‘Change the environment, not the person’

Diametrically opposed to the ‘dangers of ICT’ view (e.g., Howlin, 1998; Latash, 1998) is one that advocates the benefits of ICT for individuals with autism. So, as a contrast to interventions, the crux of this argument is that ICT can be used to change the world to suit the individual with autism rather than try to change the person with autism to fit into the world. Indeed, humans have always used technology (in its broadest sense) to augment their own cognition (Clark, 2003) and I argue that individuals with autism can similarly use ICT for their own needs.

I will now describe how different psychological theories have been used to both develop VE and understand VE use in autism. I start with the theories that have been the most influential to date, and end with those I believe hold future promise.

Theory of mind and the enactive mind

After the film show, Jim Blinn, who’s one of the pioneers in this field, came running up to me and said, ‘John, I have to ask you a question.’ And I thought, ‘God, I don’t know anything about these algorithms; I know he’s going to ask me about the shadow algorithms or something like that.’ And he asked me, ‘John, was the parent lamp a mother or a father?’

The above quote is from an interview (McCracken, 1990, p. 1) of John Lasseter, founder of Pixar, about the impact of Luxo Jr. – the film Lasseter credits as the breakthrough for computer-animated films. Luxo Jr. involves a series of interactions between a small ‘mischievous’ lamp and a larger more ‘nurturing’ larger lamp. According to Lasseter, this breakthrough came about because people started to focus on the animation rather than the technology (indeed the angle poise lamp became the icon of Pixar). In an analogous way, I argue that psychological theory should always be at the heart of development of human–computer interactions rather than the technology itself. For example, Klin et al.’s (2003) enactive mind theory states that from infancy the autistic mind is not attuned to the social world. For example, the gaze and gaze following patterns of individuals with autism are different from typically developing individuals and notably the eye region does not capture attention as strongly in those with autism (e.g., Klin, Jones, Schultz, Volkmar, & Cohen, 2002).

Klin (2000) began to construct his arguments from a social attribution task in which individuals with and without autism were shown Heider and Simmel’s (1944) silent animation of geometric objects interacting with each other. The participants with autism tended to describe the animation in mainly geometric terms: for example, ‘The big triangle went into the rectangle. There were a small triangle and a circle. The big triangle went out. The shapes bounce off each other. The small circle went inside the rectangle. The big triangle was in the box with the circle. The small triangle and the circle went around each other a few times . . . ’ (Klin, 2000, p. 840). Whereas the typically developing individuals searched for social meaning: ‘What happened was that the larger triangle – which was like a bigger kid or bully – had isolated himself from everything else until two new kids come along and the little one was a bit more shy, scared, and the smaller triangle more like stood up for himself and protected the little one . . . ’ (Klin, 2000, p. 840. Words in italic represent social attribution, from Klin, 2000).

So, in stark contrast to the autistic mind, the typically developing individuals searched for social meaning: ‘What happened was that the larger triangle – which was like a bigger kid or bully – had isolated himself from everything else until two new kids come along and the little one was a bit more shy, scared, and the smaller triangle more like stood up for himself and protected the little one . . . ’ (Klin, 2000, p. 840. Words in italic represent social attribution, from Klin, 2000).

Joint attention

ECHOES (e.g., Porayska-Pomsta et al., 2012) was designed to provide children with and without autism with learning environments in which joint attention
could be experienced and practised with a virtual character via a large touch screen. Briefly, joint attention is the ability to coordinate attention between oneself, object(s) in the environment and another person (e.g., Farrant, Maybery, & Fletcher, 2011). A classic exemplar of joint attention is proto-imperative pointing, in which the individual points out something in the environment to share with another person. This ability is thought to be unique to humans (e.g., chimpanzees can ‘point to want’, but not ‘point to show’) and its importance is so great that the development of language and social cognition (i.e., theory of mind) has been argued to be its post-cursor (e.g., Tomasello, 2008).

With respect to autism, joint attention is thought to be disrupted and, hence, leads to downstream problems in development (Charman, 2003). Indeed, there is evidence that the ability to engage in joint attention predicts subsequent language development in children with autism (McDuffie, Yoder, & Stone, 2006). The premise behind ECHOES was that if children with autism have an affinity with ICT, and they could be provided with the experience of joint attention in VE, then perhaps their joint attention ability could be improved.

With this aim in mind, Alcorn et al.’s (2011) virtual character (Paul) asked children to help him collect flowers. In this prototype learning environment, Paul indicated which one of the three flowers in the foreground he wanted, either by gazing or pointing to the flower. In addition to these gaze indicators or gestures, Paul either first ‘looked out of the screen’ (as if making eye contact with the child) or made no eye contact with the child.3

Children with autism were found to be fastest in identifying Paul’s flower selection only when his gaze or point to the flower was preceded by a ‘mutual gaze’ (Alcorn et al., 2011). Although preliminary, a rich interpretation is that children with autism were treating Paul differently depending on the context, and perhaps they were fastest when responding because they perceived the ‘mutual gaze’ as a communicative act. However, because this was an initial study no typically developing children were included to see if they had a similar pattern of responding. Another caveat is that this was not true mutual gaze (cp. Pellicano & Macrae, 2009). However, the most up-to-date ECHOES system has the potential to monitor the participant’s gaze upon the screen (via a webcam eye tracker) and so ‘knows’ where the child is looking (see Kim & Mundy, 2012, for a system of this type). In turn, the system could conceivably tell Paul where to find and ‘look’ at the participant, because it knows the participant’s location from a movement sensor. With both the eye tracking and movement sensor systems functioning, a better operationalization of mutual gaze (than anything else to date) is possible (cp. Pellicano & Macrae, 2009). Further work is clearly needed to replicate and extend these findings, but the early data are promising in that children with autism responded to Paul and did not treat him as if he were a bunch of pixels; for example, some of the children greeted and said goodbye to Paul. So, some children with autism appear to go beyond the material evidence and respond in some way to the animacy of Paul. This suggests that, while mere physical movements and collisions are insufficient to evoke animacy representations in children with ASD (Klin, 2000, above), when they interact with the richer contents of VE, greater potential is illuminated.

Intriguingly, despite the dominance of tests of false belief in the developmental literature, as the test of theory of mind (Low & Perner, 2012), there has been little evidence of this paradigm being used as a framework for VE for individuals with autism. This may stem from both engineering and theoretical considerations. First, it might be beyond our current technology to create such complex process needed for analogues of unexpected location change or unexpected context tests. Second, when researchers have tried to formally teach children with autism to pass tests of false belief it seems that the children ‘may have learned to pass the tasks rather than understanding the concepts underlying the rules’ (Hadwin, BaronCohen, Howlin, & Hill, 1996, p. 359). So, theoretically, it might be more sensible to create VE for children with autism based on the developmental precursors to false belief understanding than on the task itself.

Referential communication

The key difference between collaborative virtual environments (CVEs) and most other VEs is that collaborative VEs require two participants rather than a single user. Here, the onus is on both participants to work together. As well as conceptual problems of understanding others’ minds (theory of mind), individuals
with autism also have perceptual problems of coordinating visual perspectives (e.g., Baron-Cohen, 1989). To this end, the COSPATIAL project (e.g., Parsons et al., 2011) located pairs of children in different physical spaces from each other. However, the children had to collaborate with each other in a shared virtual space. So, for example, one child with autism would sit in front of a computer screen depicting a 3D room. While in a separate physical room, another child with autism would be in front of a computer screen depicting the same virtual room. These children then had to work in collaboration to build a tower using coloured blocks to match a pattern. To do this, the children had to select jointly a block with the colour combination that suited both their needs; that is, they had to verbally communicate (speaking via a headset) their individual needs and negotiate with someone else for mutual benefit. Early pilot data (Garib-Penna & Parsons, 2011; Parsons et al., 2011) showed that although the autistic dyads found collaborating difficult, they nevertheless could succeed with help from a facilitator.

Although COSPATIAL is the first VE to have two participants with autism working together, the theoretical basis for CVEs has a long history in developmental psychology in referential communication (e.g., Piaget, 1926; Krauss & Glucksberg, 1969; Donaldson, 1978; Robinson & Whittaker, 1987) and peer collaboration (e.g., Anderson et al., 2000). These theories, along with visual perspective role-taking and cognitive behaviour therapy, helped inform the design of COSPATIAL, thereby directly targeting aspects of autism that are known to be problematic (Garib-Penna & Parsons, 2011; Parsons et al., 2011).

**Executive function**

**Multitasking**

Multitasking is an EF, which is itself an umbrella term for all those faculties needed to work in a motivated fashion, towards a goal that may not be reached instantly (Gillberg & Coleman, 2000). These functions include planning, cognitive flexibility, generativity, inhibition, self-monitoring, etc. (see Hill, 2004, for a review of EF in autism). Rajendran et al. (2011) used the virtual errands task (VET: McGeorge et al., 2001) to look more closely at multitasking (the ability to interleave several tasks in a limited time period) in autism. The VET is a 3D representation of a building in which participants navigate along corridors and up and down stairwells using a mouse. Rajendran and colleagues used the VET to investigate if whereabout tasks performance might break down; for example, planning, memory, etc. Using a VE allowed (1) the participant to be the actual agent who has to complete the tasks (something that is hard to control ethically and experimentally in everyday life, see Shallice & Burgess, 1991, for the pitfalls of studying multitasking in the real world); (2) performance could be recorded for later detailed coding. So, arguably the VE provided both a degree of ecological validity and opportunities for fine-grained performance evaluation. Rajendran et al. (2011) found that planning inflexibility, inhibition and difficulties with prospective memory (remembering to carry out intentions) may lie behind multitasking difficulties.

**Behavioural regulation/inhibition**

As well as investigating intentionalty, a serendipitous outcome of ECHOES (Alcorn et al., 2011) was that the temporal arrangement of ECHOES seemed to help children regulate their behaviour. Specifically, in order to ‘know’ which of the three flowers Paul selected, the participant had to wait until Paul made a selection. At the beginning of experiment, some children made a flower selection without waiting for Paul’s indication. However, by the end of the experiment these children had seemingly learned to wait before reacting to Paul. In parallel to this, these children needed less physical support (gentle holding of the shoulders by the teacher) and showed less stereotypical rocking (motoric EF) by the end of experiment. So, subjectively it seemed that ECHOES had the potential to help motor self-regulation (see Sayers, Oliver, Ruddick, & Wallis, 2011, for an overview of stereotypic movement in autism). However, this needs to be corroborated by comparing the children’s baseline rocking in the classroom to that within ECHOES. This potential for motor self-regulation is made possible because movement is integral to ECHOES: that is, the child has to move to touch the screen, to act in space. So, arguably, ECHOES and other non-desk-based technologies intrinsically accommodate the motoric aspects of autism. Indeed, myself and others have argued that autism needs to be viewed in its totality as both a movement and social cognitive disorder (Rajendran &
Mitchell, 2007; Mostofsky & Ewen, 2011) – and movement-based VEs have the potential to accommodate this.

Self-monitoring

As well as being someone to interact with, a virtual character can also provide self-monitoring information. So, one useful intervention would be to provide feedback (self-monitoring) about where to look on someone’s face because eye gaze monitoring is problematic in autism (e.g., Klin et al., 2003; Norbury et al., 2009). This would be tricky to do in real life, but relatively simple in a VE. Grynszpan et al. (2012) ingeniously used a remote eye tracker to monitor gaze towards a virtual character. Using the eye tracking information, Grynszpan et al. created a ‘lens’ over the character image. If one looked through the lens, the image was clear, but outside the lens the image was opaque. During the pre- and post-test stages the lens was unconstrained (so participants could use their gaze to control the lens and look anywhere they wished). However, during the intervention the lens was constrained to the virtual character’s face. So, at this stage the participants could learn that the lens helps see the virtual character more clearly. This learning would only occur, however, if the participants were able to self-monitor, to have essentially a meta-awareness of their own eye gaze control. Intriguingly, despite the intervention, the ASD group showed a lack of awareness of being able to control the lens. Despite this initial null finding, the authors suggest that their paradigm hold promise for helping with improving self-monitoring skills in ASD.

WCC theory

Reduced generalization

No VE for individuals with autism has been yet designed using WCC theory as its principle guiding framework. However, one of the theories that came in its wake – reduced generalization theory (Plaisted, 2001) – offers insight into issues of transferring learning from a VE to the real world. This vexed issue of generalizing learning is arguably the ‘holy grail’ for any autism intervention because it has proved notoriously difficult to illicit and detect behaviour change from one context to another (Spence & Thurm, 2010). For example, from a learning environment to the real world, from the classroom to the home, etc., this difficulty is exacerbated by the plethora of unsubstantiated interventions (Dingfelder & Mandell, 2011) and the desperation of parents in seeking treatments for their children (Sharpe & Lee Baker, 2007).

Even when studies are well resourced and well specified, detecting change in such a heterogeneous condition as autism is difficult. For example (in a non-VE study), using large-scale randomized control trials, and targeting specific behaviours (i.e., joint attention), Green et al. (2010) changed behaviours at the proximal level by increasing parental interactions with their child, but this did not influence the more distal level of reducing autism symptoms. Green et al. argue that this difficulty in generalization is due to the attenuation of the effect of the intervention. So, the intervention becomes less effective (see Rao, Beidel, & Murray, 2008, for the effectiveness versus efficacy distinction) as you move from the actual point of intervention to more ‘downstream’ measures.

With the development of VE, researchers were quick to argue for the benefits of using VEs for generalizing learning from the VE to everyday life for individuals with autism (Goodwin, 2008; see Parsons & Cobb, 2011, for a recent review; Trepagnier, 1999; Parsons & Mitchell, 2002). The AS Interactive Project (e.g., Parsons et al., 2006) was a pioneering study which provided individuals with autism social skills training in queuing (how typically British) and finding a seat on a bus or in a café. Such skills may require some level of social judgment for typically developing people, but should not prove overly difficult. For individuals with autism, however, the subtleties of unwritten rules and social etiquette can be extremely anxiety provoking and debilitating. So, a VE offers a chance to practise such social skills within a managed situation. Additionally, guidance and explanations can be provided by a facilitator, and the tasks can be practised repeatedly, without fear of making a mistake or being rebuked. Additionally, the level of complexity (more people, more noise, etc.) can also be adjusted – and, so, when ready the individual can try out these skills in the real world. Although Mitchell, Parsons, and Leonard (2007) did not go as far testing learning in everyday life, they did find that behaviour learned in the VE did generalize to making judgments about where one would sit in a photograph of a bus.
I argue here that VE can be exploited to test a theoretically based (but as yet empirically untested) way to more easily generalize behaviour from the virtual learning environment to the real world. Plaisted’s (2001) reduced generalization theory not only offers an explanation as to why generalizing learning is so difficult for individuals with autism, but also a potential solution. As an alternative to WCC theory (Frith, 1989), Plaisted argues that perceptual processing in autism can be explained as reduced processing of the similarities that are held between stimuli and situations. That is, individuals with autism have narrower concepts and sharper and more clearly delineated perceptual boundaries. So, a typically developing child who learns something in the classroom should theoretically have no problem in generalizing this to his home (because he easily processes the similarities between the classroom and home). In contrast, a child with autism may find this more difficult because his processing style only allows for generalization to take place in a more perceptually gradual and graded way; for example, by making the classroom more home-like and the home more classroom-like. VE seems ideally suited to this task because the transition from one environment to another can be directly controlled. That is, by gradually changing the learning environment (Parsons & Mitchell, 2002) to make it as perceptually similar as possible to the new environment. So, to facilitate generalization, practitioners could alter the learning environment by progressing through series of intermediary ‘graded’ stages, by changing colours, textures, luminance, etc.

Enhanced perceptual processing

Where does reality lie? In the greatest enchantment you have ever experienced. (Hugo von Hofmannsthal, as cited in Curry, 2012, p. 76)

As a more radical variant of WCC theory, Mottron and colleagues argue that perception in autism is not only different, but is also superior (enhanced perceptual processing, Mottron et al., 2006). Any such qualitative difference between people with and without autism seems likely to be starker if one’s senses become totally immersed. For example, in an immersive virtual reality (IVR), the individual wears device(s) that give the wearer the sense of being in another reality (see Slater, 2009, for a review of IVR). Indeed, Slater (2009) argues that the best IVR systems provide visual, auditory and haptic input. In contrast, a VE is 3D representation on a two-dimensional display and – according to Slater – does not have the same sense of place of plausibility illusion (of really being there) of IVRs. Most autism research to date has been with VEs, but a few have investigated IVRs. Wallace et al. (2010) were the first to investigate how children with autism perceived an IVR relative to typically developing children. Immersing the children in a residential street, a school playground or school corridor allowed the researchers to record the children’s perceptual and social responses. Wallace et al. found that, although the children were passive observers, they showed similar levels of place of plausibility illusion (of really being there) as the typically developing comparison in children. More recently, Mundy and colleagues have used IVR to look at social attention (Jarrold et al., 2011). For example, Jarrold et al. (2011) investigated children’s scanning patterns when they were immersed in a virtual classroom and found that children with autism showed less attention to virtual peers furthest from the central character than IQ-matched controls. Jarrold et al. also found that attention to an array of virtual peers distinguished the autism group from the control group more powerfully than attention to non-social objects (with the autism group showing decreased attention to the virtual peers). Interestingly, attention to both social and non-social virtual characters/objects was associated with academic achievement, above and beyond variance associated with IQ in the autism sample (P.C. Mundy, personal communication, August 1, 2012).

Future research questions, in this area, might be to investigate if the perceived levels of immersion in IVRs interact with the perceptual superiority (Mottron et al., 2006) that individuals with autism might have. Additionally, IVR arguably holds still greater promise because it has both sensory motor and social cognition qualities (e.g., through interacting with virtual characters) – both of which are affected in autism (Rajendran & Mitchell, 2007; Sayers, Oliver, Ruddick, & Wallis, 2011; Wallace et al., 2010).

The question of reality is, however, not only just about the physical environment, but also about the nature of the virtual people populating that environment. This, too, is open for insight via studies with individuals with autism. We already know that
individuals with autism are less susceptible to prior (top-down) knowledge interfering with their object visual perception than typically developing individuals (Ropar & Mitchell, 2002). What we do not know is if there is a point at which they might baulk at virtual characters for being ‘too human’.

The uncanny valley phenomenon (Mori, 1970) occurs when people show a disdain for representations of humans that are too close to human form for comfort (see also Parsons, 2007; Gray & Wegner, 2012; Pollick, 2010). So, people generally have no aversion to robots and other human analogues provided they are not too human. As they become perceptually more human, however, a point is reached (the uncanny valley) when these creations are no longer accepted. To date, the virtual humans in autism research have rendered virtual characters from stylized cartoon forms (e.g., Porayska-Pomsta et al., 2012) to more human-like representations (e.g., Ehrlich & Miller, 2009; Cheng & Ye, 2010; Wallace et al., 2010). The style in which to depict virtual humans can be a question of resources as much as of design. However, active decisions have also been made in not trying for too much human realism (e.g., Porayska-Pomsta et al., 2012). Arguably, if the content is right, then striving for realism may not be as important as other factors. In fact, it would be interesting to investigate whether individuals with autism prefer virtual humans who are more or less realistic. However, this question might be moot provided the participant imbues the character with mental life – that is, has prior (top-down) expectations of the character’s behaviour (Alcorn et al., 2011; Porayska-Pomsta et al., 2012).

Pollick (2010) states that the uncanny valley is ready for renewed investigation. He argues that this phenomenon might result from issues of physical perception, biological motion perception and/or social cognition. Autism is affected in all these areas and so autism research into this phenomenon would be insightful into the boundaries of what humans in general perceive as being ‘real’ or not.

**Embodied cognition**

An increasingly important question in psychology is how the body shapes the mind (e.g., Clark, 1999; Wilson, 2002; Borghi & Cimatti, 2010), that is, embodied cognition. The Cartesian view of the separateness of mind and body has been a dichotomy that has arguably held back our understanding (Glenberg, 2006). Nowhere is this seen more starkly than in autism, which is either conceived as a movement disorder (Mostofsky & Ewen, 2011) or a social–emotion disorder, rather than in its entirety (Rajendran & Mitchell, 2007; Mostofsky & Ewen, 2011). Specifically, the movement aspects (e.g., Sayers et al., 2011) seem divorced from the emotion and cognitive aspects (Rajendran & Mitchell, 2007), when in fact they should arguably be considered in a more unified way. Imitation researchers have, however, been trying to link these aspects of development (see Rogers & Williams, 2006, for a review of the imitation research field). This is because imitation requires both representational aspects (i.e., relating self to other for copying another’s actions) and motoric aspects (i.e., actually executing a movement that one wishes to copy) (Stieglitz Ham et al., 2011). For example, we can use gesture imitation to find out exactly at what point the individual has an imitative problem. So, if someone with autism fails to imitate a gesture, we can now pinpoint whether it is because he cannot adequately represent that gesture in his mind, or because he fails to execute the gesture movement (e.g., Stieglitz Ham et al., 2011).

In tandem with research into imitation in autism, there has been a recent resurgence in the study of embodied cognition (Goldman & de Vignemont, 2009; Porayska-Pomsta, Bernardini, & Rajendran, 2009; Borghi & Cimatti, 2010; Glenberg, 2010). For example, Tversky and Hard (2009) found that when typically neurotypical adults saw a photograph of a person reaching for an object it helped improve their own perspective taking of the scene (indeed the mere presence of the person helped when compared with a photo with no one present). Additionally, Cook and Tanenhaus (2009) found that a speaker’s hand gestures (but not his/her speech) encoded information about solving a planning task (Tower of Hanoi) to a listener. There are also examples in the developmental literature: for instance, babies who crawled through a space found hidden objects more than babies who were carried (Benson & Uzgiris, 1985). In another study, children who walked through an imagined space were better able to state the locations of objects in the real space than are those who sat still while imaging the space (Rieser, Garing, & Young, 1994; for reviews, see Glenberg, 2010 and Lillard, 2007, Chapter 2).
So, both the fields of imitation and embodied cognition allude to the importance of being able to move in space for cognitive development. This was taken into consideration in the design of systems such as ECHOES (e.g., Porayska-Pomsta et al., 2012) in which the child has to move to interact with the virtual character (as opposed to sitting with a console in hand). The capacity to move may be highly functional for individuals with autism. Indeed, the stereotypic movement of some children with autism may help regulate their levels of arousal (McDonnell, 2010). Hence, the children who regulated their own stereotypic rocking (see above) took advantage of this aspect of ECHOES. Additionally, the relation between movement and cognition might be more readily studied in such VE systems – because these systems not only provide the opportunity for the child to move, but also can measure the movement non-invasively.

Concluding thoughts

Each of the theories I have discussed has something to offer VE research and autism. They can provide a framework, insight, direction and inspiration – and arguably, without them, the technologies might not fulfil their full potential. This is because each theory provides a different a perspective on human psychology. So, VE research in autism includes some of the most quintessential aspects of the human experience: from understanding others’ intentions (theory of mind) to managing everyday tasks like shopping (multitasking). Part of being human is also how we have always used technology to augment our own minds, and this area of research also provides insight on this process.

Psychological theory often has relevance to central research questions, thereby potentially informing the initial design stage, through evaluation and back again to theory. This process is naturally iterative and one that I believe will enhance our understanding of autism. However, the pay-offs are potentially greater still. This is because autism can be considered a disorder of social cognition, of cognition, of emotion, of perception and of movement, and as such, it provides a unique window onto human–computer interaction. This ‘finessing’ of the human condition helps illuminate these different aspects of human–computer interaction – and through careful study we can create better technologies for everyone.

Despite its importance, theory is only one of many other considerations that will guide the development of technologies; others include educational objectives, user-centred design processes, usability, affordances of the technology, and so forth. These considerations all stem from the fundamental aims of the research – which can range from interventions (e.g., teaching social skills) to basic research (e.g., investigating the process of multitasking). So, theory is just one aspect in the development and evaluation process of VE for individuals with autism – and crucially the point at which theory is enacted will very much depend on these initial research aims.

Finally, as we enter an era in which games consoles and keyboards are being replaced by our own movements, the interface is likely to change (perhaps to include interactions with virtual humans) – and our minds will evolve as they have arguably always done. This brave new world also has the potential to be a more accepting one, by including previous excluded members of society – like individuals with autism. The fundamental premise of the discipline of human factors and ergonomics – in the design process – is to fit the system to the person, not the other way around (Grandjean, 1980). So, we might consider using technology in analogous way to change the environment and not the person.

Acknowledgements

I would like to thank Katherine Myant, Kevin Durkin and the two reviewers for their help and invaluable comments in developing this paper. I also thank Neil Hendry for helping with the literature review and with editing.

Notes

1See Gernsbacher, Sauer, Geye, Schweigert, and Goldsmith (2008) for why autistic is a respectful term.
2See Mundy and Newell (2007) for a comprehensive description of joint attention.
3Previous work has shown that typically developing 1-year-olds know that when an adult makes direct eye contact, this is perceived by the child as a communicative act (Behne, Carpenter, & Tomasello, 2005).
4Accuracy was equally high in all conditions.

References


Pollick, F. E. (2010). In search of the uncanny valley. In P. Daras & O. Mayora (Eds.), *User centric media* (pp. 69–78). Berlin Heidelberg: Springer.


