Central Lancashire Online Knowledge (CLoK)

It is advisable to refer to the publisher's version if you intend to cite from the work.
https://doi.org/10.1002/aur.1352

For information about Research at UCLan please go to http://www.uclan.ac.uk/research/

Running head: EYE MOVEMENTS IN AUTISM SPECTRUM DISORDERS

Cognitive Perspective-Taking during Scene Perception in Autism Spectrum Disorder: Evidence

from Eye Movements

Sheena K. Au-Yeung[1], Johanna K. Kaakinen[2], Valerie Benson[1]

[1]University of Southampton

[2]University of Turku

Address:

School of Psychology,

Shackleton Building,

University of Southampton,

SO171BJ

Tel: 02380594596

Corresponding Author: Valerie Benson

e-mail: V.Benson@soton.ac.uk

Lay Abstract

The present study investigates the differences between adults with Autism Spectrum Disorders (ASD) and typically developing (TD) adults in the way they viewed scenes when they are required to take on the perspective of another person. Participants had their eye movements recorded while they looked at some pictures of household scenes after being instructed to "imagine that you are a burglar" or "imagine that you are a repairman". They also completed a non-perspective-taking control task in which they were asked to "look for the valuable items" or "look for the features of the house that need fixing" in another set of household scenes. Both groups showed a preference to look faster and for longer at task relevant items for the burglar task and the non-perspective-taking equivalent "look for the valuable items" task. However, eye movements showed that participants with ASD were initially slower than TD participants to look to the task relevant parts of a scene during the repairman task and the non-perspective-taking equivalent "look for the features of the house that need fixing" task. We discuss how the ambiguity of the repairman and its non-perspective equivalent task could be driving the difficulty experienced by participants with ASD.

Abstract

The present study examined how eye movements during scene viewing are modulated by perspective-taking in both adults with Autism Spectrum Disorders (ASD) and typically developing (TD) adults. In the current study, participants viewed house scenes with either non-perspective-taking (look for valuable items/features of the house that need fixing) or perspective-taking instructions (imagine that you are a burglar/repairman) while their eye movements were recorded. The eye movement measures revealed that for the "look for the valuable items" and burglar perspective task the ASD group showed typical relevance effects (the preference to look at schema-relevant compared to schema-irrelevant targets) in their eye movements. However, we found subtle processing differences between the groups that were related to initial orienting to and processing of schema-relevant items for the "look for the features that need fixing" and the repairman perspective-taking task. There was an absence of a relevance effect in the early eye movement measures for the ASD group for the repairman perspective and its non-perspective-taking equivalent instruction showing that the identification of items relevant to those schemas was more difficult for the ASD group. The present findings suggest that resolving ambiguity may be a defining feature of complex information processing deficits in ASD (Minshew & Goldstein, 1998).

Key words: Autism, Asperger's Syndrome, Theory of Mind, Information Processing, Neuropsychology, Cognition.

Introduction

Autism Spectrum Disorder (ASD) is a lifelong neurodevelopmental condition diagnosed by abnormalities in social interaction, communication, and repetitive, restricted and stereotyped patterns of behaviours and interests (APA, 2000). One influential cognitive explanation of ASD is the Theory-of-Mind (ToM) Deficit Hypothesis, which proposes that the inability to infer mental states of others underlies the social communicative deficits in ASD (Baron-Cohen, 2001).

Recent theoretical development has extended the original ToM deficit hypothesis by proposing that the social and communicative deficits observed in ASD result from a delay in the development of an empathizing system, whereas intact or superior skills in systemizing are thought to offer an account for non-social aspects of ASD such as repetitive behavior or narrow interests. In this two-factor Empathizing-Systemizing theory (Baron-Cohen, 2009), the systemizing dimension refers to the drive to construct and analyze systems and the empathizing dimension refers to the ability to identify mental states in others and to produce appropriate emotional responses. It is argued that the dissociation between these two dimensions can act as a reliable indicator of whether or not someone has ASD. Furthermore the theory would suggest that it would be impairments in the empathizing dimension that would lead to a reduced or absent ability for individuals with ASD to take on the perspective of another person.

An alternative contemporary theory of ASD, namely the disordered complex information processing theory (Minshew, Williams, & McFadden, 2008) proposes that it is a reduction in the capacity to process complex information, across cognitive domains, which underpins ASD. Complex information processing tasks require one or more of the following: integration of multiple features, speed of processing, processing of large amounts of information and processing of novel stimuli or information (Minshew et al., 2008). Minshew, Goldstein, and Siegel (1997) found that individuals with ASD performed at a reduced level compared to typically developing

(TD) individuals in a large battery of tasks designed to test higher-order cognitive processing, for example 'concept formation aspects of abstraction', whereas performance was intact or enhanced in the same ASD sample for tasks that tested basic or mechanical abilities, for example 'rule learning aspects of abstraction'.

In summary, despite the different use of terminology in these two different theoretical contexts, both theories (Baron-Cohen, 2009; Minshew et al., 2008) predict that individuals with ASD should have problems dealing with tasks that are unsystematic or complex. Perspective-taking, an aspect of empathizing, is a complex task that requires for example; activating and maintaining knowledge about what other people would be interested in, and making use of various contextual cues to infer the mental state of the other person. Thus, perspective-taking can be assumed to be challenging to individuals with ASD.

Pichert and Anderson (1977) studied perspective-taking in TD participants by asking them to read a story describing the interiors of a house from either a homebuyer's or a burglar's perspective. They showed that adopting the psychological perspective of another person while reading a story resulted in increased memory for perspective relevant information. Kaakinen, Hyönä, and Viljanen (2011), examined the eye movements of TD adults during perspective-taking tasks. When TD individuals view scenes with a specific psychological perspective in mind, more total time and more fixations were made to perspective-relevant than perspective-irrelevant areas in the scenes. Furthermore, while the first fixation was more likely to land on a visually salient than a non-salient target region irrespective of the relevance of a target, this saliency effect is quickly overridden by task instruction. In the present study, we used a modified version of the Kaakinen et al. (2011) paradigm to investigate perspective-taking in ASD. This is an important extension to previous work as it addresses whether deficits in perspective-taking, which is an aspect of ToM, are related to ASD

Eye movements are known to reflect the moment-to-moment cognitive processes during the time-course of a task (Rayner, 2009), and as such they can reveal subtle processing differences between TD and ASD individuals that are not available from other behavioral measures such as response time and accuracy. A ToM study that recorded eye movements and verbal responses (Senju, Southgate, White, & Frith, 2009) found that adults with Asperger's Syndrome were able to pass standard verbal first-order and second-order belief attribution tasks, but, unlike the TD group, their eye movements revealed no spontaneous orienting to the information. One of our studies presented pairs of scenes to investigate processing for a simple "spot the difference" task, and a more complex "which one's weird" task. Eye-movement sequence analyses (Au-Yeung, Benson, Castelhano, & Rayner, 2011) that took into account spatial location and the sequential and temporal nature of the eye fixations revealed that the eye movement patterns differed between TD and ASD participants exclusively for the complex task. A regions-of-interest analysis (Benson, Castelhano, Au-Yeung, & Rayner, 2012) showed that ASD participants took longer to begin fixating the "weird" target region, and that they did not immediately grasp what was weird as soon as they fixated it, despite performing at ceiling in terms of the accuracy measure. Hence, while TD individuals are known to modulate their eye movements when instructed to make a higher-level inference, this modulation is not always observed in ASD (Benson, Piper, & Fletcher-Watson, 2009).

The current study aimed to investigate whether individuals with ASD can take the psychological perspective of another person, and recorded eye movements to examine whether there were processing differences between ASD and TD individuals during non-perspective-taking and perspective-taking tasks. Perspective-taking tasks required participants to infer the category of objects that the perspective characters would be interested in, and then interpret whether information in the scene (pictures of the inside rooms and the outside of houses) belonged to that category (Kaakinen et al., 2011). It was predicted that TD participants with a

burglar schema in mind would be interested in valuable items in the house, whereas things that needed fixing in the house would be of interest to the participants with a repairman schema in mind. Consequently, participants should attend faster to, and remain longer looking at the schema-relevant items compared to the irrelevant items. We termed the difference between schema-relevant and irrelevant items in the dependent eye movement measures as the *relevance effect*. We also devised two non-perspective-taking tasks in which participants were explicitly told to look for a certain category of objects in the scene. The non-perspective-taking tasks were designed to direct participants' attention to the same targets as for the perspective-taking tasks, but without the need to adopt the psychological perspective of another person. Theory-of-mind would predict ASD individuals to be unable to take on the perspective of another because of an inability to empathize, which would be reflected in the eye movement records as the lack of a relevance effect for the perspective-taking task. In the non-perspective-taking task, however, individuals with ASD are expected to demonstrate the relevance effect. In contrast, the complex information processing theory would predict that participants with ASD would be unable to take on the perspective of another person because perspective-taking is a complex task that involves integrating higher level cognitive information, compared to the non-perspective-taking task. In that case ASD individuals may not show any relevance effect in the perspective-taking task especially for the early processing measures – relevance effects may, however, show up in the later processing measures indicating that individuals with ASD need more time to complete the complex task. Similar to predictions for the ToM theory, the disordered complex information processing theory (Minshew & Goldstein, 1998) predicts that both groups (TD and ASD) should show the relevance effect in the non-perspective-taking task.

Method

Participants

The participants were 17 TD men ($n = 13$) and women ($n = 4$) and 18 men ($n = 15$) and women ($n = 3$) with ASD. Typically developing participants were recruited from the local community. The participants with ASD were recruited from the Southampton Adult Asperger's Society, the Hampshire Autistic Society, the Autism Diagnostic and Research Centre, the National Autistic Society, and the Children on the Autistic Spectrum Parents' Association. Prior to the study, participants with ASD were clinically diagnosed in the UK under the criteria of International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (1992) for an ASD. Diagnostic reports confirmed that participants with ASD were primarily diagnosed using standard diagnostic instruments including the Adult Asperger Assessment (Baron-Cohen, Wheelwright, Robinson, & Woodbury-Smith, 2005), Autism diagnostic observation schedule (Lord, Rutter, DiLavore, & Risi, 2001), and/or the Autism diagnostic interview–Revised (Lord, Rutter, & Le Couteur, 1994).

All participants completed the Autism-Spectrum Quotient (AQ: Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), and the Wechsler Abbreviated Scale of Intelligence (The Psychological Corporation, 1999), which confirmed that the two groups were matched on all IQ-subsets. The ASD group scored disproportionately higher on the AQ than the TD group. Comparing ASD and TD adults who fall within the normal IQ range allows the identification of cognitive features that are unique to ASD and that are not attributable to intellectual disability (Minshew & Williams, 2008). Participants with missing data in any of the experimental conditions were excluded. This included a woman with ASD due to calibration error and a man with ASD due to computer crash. Other participants, including a man with ASD with a history of head injury, a TD man scoring over the autism cut-off and two women (one TD and one with ASD), who scored lower than 80 on all IQ-subsets were also excluded from the analysis. The characteristics for the final participants sample (TD = 15, ASD = 14) are summarized in Table 1.

[INSERT TABLE 1 HERE]

Materials

The stimuli were color photographs of two houses. For each house there was a set of 8 pictures (See Figure 1 for an example and Supplementary Online Materials for all the pictures). The pictures (1024 x 768 pixels) were taken using a Canon Powershot A640 digital camera and digitally manipulated. Each picture contained equal numbers of relevant and irrelevant target items (one, two, or three of each type) and there were no significant differences in the size of the two item types (valuable items: $M = 39276$ pixels, features needed that fixing: $M = 32367$ pixels, $p = .446$). For the burglar schema, the valuable items were the relevant targets and the features of the house that needed fixing were the irrelevant targets. This was reversed for the repairman schema. The pictures within each set were presented in a fixed order for all participants, always starting with picture of the exterior of the house, followed by the interior pictures.

[INSERT FIGURE 1 HERE]

A pilot study was conducted to verify the appropriateness of the items for each perspective. Six TD individuals who did not participate in the actual experiment were given colored prints of the stimuli on paper and asked to circle the items that were relevant for a burglar or a repairman. The order of which perspective was presented first was counterbalanced across participants. The rating study verified that the relevant items on average were identified by participants 93% of the time, and there was no significant difference between the two perspectives (burglar: 97%, $SD = 19$, repairman: 89%, $SD = 9$), $t(21.5) = 1.59$, $p = .127$.

Apparatus

Participants viewed the stimuli binocularly at a viewing distance of 70 cm on a 19 inch monitor with a screen resolution of 1024 by 768 pixels and a refresh rate of 75 Hz. Eye movements were recorded monocularly for the right eye using an Eyelink 1000 eye-tracker (SR

Research Ltd, Osgoode, Canada) at a sampling rate of 1000 Hz. A chin rest and a forehead support were used to maintain participants' head position. Participants were calibrated using a nine-point matrix where they were required to fixate each calibration point in a random sequence. This was repeated to validate that each fixation was within 0.5 degrees of visual angle of corresponding calibration points.

Design

The experiment was a mixed design with one between-participants factor: *Group* (ASD vs. TD) and two within-participant factors: *Item Type* (valuable items vs. features that need fixing) and *Schema* (burglar vs. repairman for the perspective-taking task; 'Look for the valuable items' vs 'look for the features that need fixing' for the non-perspective-taking task).

Procedure

Ethics and Research Governance approval was obtained from the University of Southampton. Participants gave written consent and took part in all schema conditions following verbal task instruction from the experimenter. A calibration procedure preceded each set of picture trials. For the two non-perspective-taking tasks participants were instructed to "*look at the items of the house that are valuable", or "look at the features of the house that need fixing".* For the two perspective-taking tasks, participants were instructed to *"look at the pictures and imagine that you are a burglar"* or *"look at the pictures and imagine that you are a repairman".* Prior to the presentation of each picture, participants had to fixate a central dot. Recalibration was carried out if participant's point of gaze no longer matched the location of the fixation dot.

Participants viewed each picture set twice; i.e., if a participant started with the non-perspective-taking task they would view the pictures with one of the instructions (e.g., look for valuable items) and then be asked to view the same picture set with the other instruction (look for features that need fixing). A different set of pictures was presented to participants for the non-

perspective-taking and perspective-taking tasks and the order of the two different picture sets and the order of the tasks were counterbalanced across participants. Each picture was presented for six seconds.

Following the eye-tracking task, all participants completed a pen and paper study. They were presented with a word list of 30 items twice and were instructed to tick the box next to a word if it referred to 'valuable items in a house' or a 'feature of a house that needed fixing'. The order in which the two instructions were given was counterbalanced across participants. The word lists contained ten valuable items, ten features that needed fixing, and ten neutral items (e.g. a shoe box). The word list study confirmed both groups of participants were equal in their ability to identify items that belonged to each category (valuable items: TD: $Mdn = 10$, ASD: $Mdn = 10$, $U(15, 14) = 98$, $p = .705$; features that need fixing: TD: $Mdn = 10$, ASD: $Mdn = 10$, $U(15, 14) = 103.5$, $p = .936$). Participants were debriefed at the end of the study.

Eye Movement Analysis

Fixations and saccades were identified using a velocity criterion of 30 degrees per second. Any fixations that spanned across screen changes or were shorter than 50 ms were discarded. Eyelink DataViewer was used to create freehand regions of interest surrounding the shape of each target. For each dependent variable involving regions of interests, extreme data points were removed if they were two or more standard deviations away from an individual's mean over an entire condition (across eight pictures) for each type of target ($< 3\%$ data).

The data were analysed separately for the non-perspective-taking and perspective-taking tasks. For the non-perspective-taking task a 2 Schema (look for the valuable items vs. look for features that need fixing) × 2 Item Type (valuable items vs. features that need fixing) × 2 Group (ASD vs. TD) repeated measures ANOVA was computed for each dependent variable. For the perspective-taking task a 2 Schema (burglar vs. repairman) × 2 Item Type (valuable items vs. features that need fixing) × 2 Group (ASD vs. TD) repeated measures ANOVA was computed for

each dependent variable. Significant interactions were followed up with pairwise t-tests; Bonferroni corrections to the p-values were applied to control for the possibility of inflated family-wise error rates.

Results

Baseline Eye Movement Measures

Baseline eye movement measures were computed across all trials and conditions for each group. There were no between-group differences in the mean total viewing time spent on the stimuli (TD = 4937 ms, ASD = 4925 ms), mean total number of fixations (TD = 18.4, ASD = 17.3), mean fixation duration (TD = 282 ms, ASD = 314 ms), first saccade latency (TD = 233 ms, ASD = 254 ms), and saccade amplitude (TD = 6.00, ASD = 5.78), all p's > .05. These findings indicate that any observed between-group differences in the regions of interest analysis in the following sections are unlikely to be the result of between-group differences in basic sampling and oculomotor control.

Global Eye Movement Measures

We analysed global eye movement measures, including, *total viewing time*, *total number of fixations*, and *mean fixation duration* within the predefined target areas (valuable items and items that needed fixing). These measures give an indication of the importance of the target in the scene related to the task at hand. Greater viewing time and a higher number of fixations means that more attention was allocated to the area and more processing of the contents was carried out (Benson et al., 2012). Descriptive statistics for the global measures are presented in Table 2.

[INSERT TABLE 2 HERE]

Total viewing time. In the *non-perspective-taking task*, there was a main effect of Schema, $F(1,27) = 15.9$, $p < .001$, $\eta_p^2 = .371$, indicating that overall, more time was spent on target regions when viewing the pictures with the 'look for valuable items' ($M = 1490.82$ ms)

than with the 'look for features that need fixing' schema ($M = 1292.58$ ms). A main effect of

Item Type, $F(1,27) = 42.95$, $p < .001$, $\eta_p^2 = .614$, showed that participants spent more time on

valuable items ($M = 1623.82$ ms) than on features that need fixing ($M = 1159.59$ ms). More

importantly, there was a Schema × Item Type interaction, $F(1,27) = 170.43$, $p < .001$, $\eta_p^2 = .863$,

indicating that more time was spent on schema-relevant than irrelevant target regions for both the

'look for valuable items' ($t(28) = 15.15$, $p < .001$, $d = 2.813$) and the 'look for features that need

fixing' schema ($t(28) = 8.91$, $p < .001$, $d = 1.655$).

Similarly, in the *perspective-taking task*, there was a main effect of Item Type, $F(1, 28) =$

6.71, $p = .015$, $\eta_p^2 = .199$, (valuable items: $M = 1430.72$ ms, features that need fixing: $M =$

1129.92 ms), and an interaction between Schema and Item Type, $F(1,28) = 112.08$, $p < .001$, η_p^2

$= .806$, indicating that more time was spent on schema-relevant than irrelevant target regions for

both the burglar ($t(28) = 9.65$, $p < .001$, $d = 1.792$) and the repairman schema ($t(28) = 7.05$, $p <$

001, $d = 1.308$).

In sum, the results of the total viewing time show a clear relevance effect for both non-

perspective-taking and perspective-taking schemas for both groups (see top left of Figure 2). The

number of fixations results mirror the results for the total viewing time (see Supplementary

Online Materials).

<div align="center">[INSERT FIGURE 2 HERE]</div>

Mean fixation duration. In the *non-perspective-taking task*, there was a Schema × Item

Type interaction, $F(1,27) = 42.16$, $p < .001$, $\eta_p^2 = .610$, indicating that participants made longer

fixations on schema-relevant than irrelevant target regions across conditions and groups for both

the 'look for valuable items' ($t(28) = 5.05$, $p < .001$, $d = 938$) and 'look for features that need

fixing' schemas ($t(28) = 6.76$, $p < .001$, $d = 1.255$). In the *perspective-taking task*, a Schema ×

Item Type interaction, $F(1,27) = 37.29$, $p < .001$, $\eta_p^2 = .580$, indicated that participants made

longer fixations on schema-relevant than irrelevant target regions for both the burglar ($t(28) =$

6.06, $p < .001$, $d = 1.125$) and the repairman schema ($t(28) = 5.27$, $p < .001$, $d = .978$). In sum, clear relevance effects for mean fixation duration were shown for both non-perspective-taking and perspective-taking schemas for both groups.

Early Eye Movement Measures

We also computed early processing measures for the predefined target areas, including *elapsed time* and *fixation count to target*, which give an indication of early orienting towards the target areas, and of how long participants spent exploring and processing other parts of the scene before a target captured the viewer's attention. *First fixation duration* gives an indication of recognition of the item as a target when it is initially attended to. Descriptive statistics for the early measures are presented in Table 3.

Elapsed time to target. In the *non-perspective-taking task*, there was a significant main effect of Item Type, $F(1,28) = 19.28$, $p < .001$, $\eta_p^2 = .417$, indicating that it took longer to look at features that need fixing ($M = 1502.60$ ms) than at valuable items ($M = 967.74$ ms). The main effect was qualified by a two-way interaction between Schema and Item Type, $F(1,27) = 79.28$, $p < .001$, $\eta_p^2 = .746$, indicating that participants landed sooner on schema-relevant than on irrelevant target regions. The three-way interaction between Schema, Item Type, and Group was also significant, $F(1,27) = 4.71$, $p = .039$, $\eta_p^2 = .149$. Pairwise comparisons revealed that the TD group were faster to fixate on schema-relevant than irrelevant target regions for the 'look for valuable items' schema ($t(14) = 7.97$, $p < .001$, $d = 2.058$), and marginally for the 'look for features that need fixing' ($t(14) = 2.35$, $p = .034$, $d = .607$), whereas the ASD group were only faster to fixate on schema-relevant than irrelevant target for the 'look for valuable items' schema ($t(13) = 4.50$, $p = .001$, $d = 1.202$) and not for the 'look for features that need fixing' schema ($t(13) = 1.53$, $p = .151$, $d = .408$).

For the *perspective-taking task*, a significant main effect of Item Type, $F(1,27) = 11.21$, $p = .002$, $\eta_p^2 = .293$, shows that it took longer to land on features that need fixing ($M = 1592.03$ ms)

than on valuable items ($M = 1170.22$ ms). A Schema × Item Type interaction $F(1,27) = 92.39$, p < .001 , $\eta_p^2 = .774$, indicated that overall participants were faster to fixate on schema-relevant than irrelevant target regions. In addition, a three-way interaction between Schema, Item Type, and Group that was approaching significance, $F(1,28) = 4.02$, $p = .055$, $\eta_p^2 = .130$, showed that elapsed time to schema-relevant and irrelevant target regions was modulated by Group. Follow-up comparisons clarified that TD participants were faster to fixate on schema-relevant than irrelevant target regions in both the burglar ($t(14) = 9.89$, $p < .001$, $d = 2.554$) and the repairman schema ($t(14) = 5.47$, $p < .001$, $d = 1.412$), whereas in the ASD group the difference between schema-relevant and irrelevant target regions was significant for the burglar ($t(13) = 4.04$, $p = .001$, $d = 1.080$) but not for the repairman schema ($t(13) = .601$, $p = .558$, $d = .161$).

As can be seen in Figure 2, the TD group showed *relevance effects* for elapsed time to target in all schema conditions, whereas the ASD group, showed *relevance effects* for the burglar and the non-perspective-taking counterpart 'look for the valuable items' schema, but not for the repairman schema and its non-perspective-taking counterpart 'look for the features that need fixing' schema (see Figure 2 Bottom Left). The *fixation counts to target results mirror* those for the elapsed time to target (see Supplementary Online Materials).

First fixation duration. In the *non-perspective-taking task*, there was a two-way interaction between Schema and Item Type, $F(1,27) = 26.52$, $p < .001$, $\eta_p^2 = .496$, suggesting that first fixation durations were longer on schema-relevant than on irrelevant target regions for both the 'look for valuable items', ($t(28) = 3.87$, $p = .001$, $d = .719$) and the 'look for the features that need fixing' schema, ($t(28) = 3.18$, $p = .004$, $d = .591$). In the *perspective-taking task*, there was a marginal main effect of Schema, $F(1,27) = 3.18$, $p = .086$, $\eta_p^2 = .105$, suggesting that first fixation durations were slightly longer when viewing the pictures with the burglar ($M = 268.02$ ms), than the repairman perspective ($M = 249.22$ ms). An interaction between Schema and Item type, $F(1,27) = 4.84$, $p = .037$, $\eta_p^2 = .152$, indicates that first fixation durations were significantly

longer on schema-relevant compared to irrelevant target regions for the burglar schema ($t(28) =$ 2.79, $p = .009$, $d = .009$), but not for the repairman schema ($t(28) = .127$, $p = .900$, $d = .024$).

The relevance effect on first fixation duration for the non-perspective-taking task was clear for both schemas across both groups, but was present only for the burglar schema in the perspective-taking task.

[INSERT TABLE 3 HERE]

Discussion

The current study explored whether or not individuals with ASD, like TD individuals (Kaakinen et al., 2011), were able to take on the perspective of another person during scene inspection. This is the first eye-tracking study to manipulate complexity in terms of non-perspective and perspective-taking tasks. Recently, Loth, Gomez, and Happé (2011) investigated the influence of event knowledge on attention and memory for context-relevant aspects of a scene and an effect of relevance was not found for either proportion of gaze time and number of fixations for both TD and ASD groups, but was observed in the average fixation duration for both groups. In the current study, a modulation of eye movements in the global measures for both the non-perspective-taking and perspective-taking schemas was consistently observed for both participant groups. This means that task-relevant elements of a scene were more likely to receive attention, were inspected for longer and were processed in more depth than the task irrelevant parts of a scene by all participants. We suspect that the differences in findings between the two studies could reflect methodological differences. Participants in the study by Loth et al. were not actively instructed to inspect the scene with the schemas in mind, and their stimuli contained a much larger array of objects with a much longer stimulus presentation time. These factors could have driven participants to explore non-context relevant aspects of the display more, therefore failing to show the robust relevance effects for all global measures observed in the current study. In contrast with the ToM prediction, our results clearly indicate that, given enough time,

participants with ASD show similar perspective-taking behaviour as TD individuals as reflected in global eye movements measures over the entire duration of the tasks. And, also in contrast with our original interpretation of disordered complex information processing theory, there were subtle processing differences for the two different perspectives.

The findings from our early eye-movement measures showed that individuals with ASD do not show initial orienting to, and processing of, target relevant items equally for the two different perspectives and across the two non-perspective-taking instructions. This suggests that complexity (Minshew et al., 2008) cannot be defined by the perspective-taking element of a task per se, but it could be accounted for by qualitative differences that might exist between the different perspectives. For example, for the burglar schema and its non-perspective-taking equivalent 'look for the valuable items' schema, both TD and ASD showed a relevance effect in all three early measures, such that they spent less time and made less fixations to other parts of the scenes before orienting to schema-relevant compared to irrelevant targets. Furthermore, they spent more time fixating a relevant compared to an irrelevant target the first time they attended to it, suggesting they are able to immediately identify the relevant item for the burglar and the 'look for the valuable items' schema. In contrast, for the perspective-taking repairman schema and its non-perspective-taking equivalent "look for the feature that needed fixing" schema, participants with ASD showed an absence of a relevance effect for elapsed time to target, and for fixation count to target.

The difference between the two perspectives could be related to the fact that 'features that need fixing' tend to be structural features of the house at the background of the scene, whereas burglar relevant items tend to be foreground objects. In general, viewers prefer to look at foreground objects over background objects (Yarbus, 1967), because it is more likely that these objects contain potentially meaningful semantic information (Henderson, Malcohm, & Schandl, 2009). Therefore, even during the 'look for the features that need fixing' task or the repairman task, the 'valuable items' will inevitably attract attention initially, purely because they are at the

foreground, and this may have contributed to a reduced relevance effect even for the TD group for the repairman perspective and the 'look for features that need fixing' schema. It could also be that individuals with ASD were slower in overcoming this default mode of initial attention by guidance of the task instructions than TD individuals. The following sections discuss two possibilities as to why this might be the case.

Pichert and Anderson (1977) suggested that some perspectives are harder to keep in mind than others due to unfamiliarity with the role. It is possible that in the present study, the novelty of the repairman and 'look for the features that needed fixing' schema made it harder to draw on prior knowledge on what targets are relevant to the tasks and therefore led to a greater exploration of irrelevant targets initially. It could also be that individuals with ASD have less experience with home repairs because more of them, in comparison to TD individuals, continue to live with caregivers who are likely to take care of household maintenance rather than living independently. This is however only a speculation, as we did not collect data regarding the living arrangements of our participants.

A more likely explanation for the lack of a relevance effect in the repairman and the 'look for the features that needed fixing' schema observed in the ASD group relates to possible ambiguities in the categorization of the relevant items in the pictures. For example, when participants were asked to imagine they were a burglar, their role was clear to them, and they were able to quickly relate to the idea that a burglar would be interested in valuable items in a house. In a previous study by Kaakinen and Hyönä (2008) the burglar perspective proved to be a relatively consistent concept within a sample of college students. Kaakinen and Hyönä asked participants to list things that would be of interest to a burglar and found that there was considerable overlap in what kind of items were listed by different participants. However, a repairman's role might be more ambiguous because it could be interpreted as someone who fixes objects other than the structural features of the house. And similarly, 'features that need fixing' could refer to object features rather than structural features of the house. This means that there are

potentially multiple ways to interpret the repairman task instructions or the items that might belong to that category.

In relation to ToM and the extended Empathizing-Systemizing theory (Baron-Cohen, 2009), the results from the current study clearly raise question to whether deficits in ASD are restricted to ToM or empathizing and also question whether a better explanation or working definition of ToM is now needed for future investigation of these concepts. Recently, Baron-Cohen (2009) suggested that the two dimensions (empathizing and systemizing) could be reduced to a single dimension defined by "the extent to which one is able to deal with degrees of unlawfulness in information" (p. 78). Lawfulness of information can be thought of as information that is rule based and which has a single predictable outcome. It is the process of systemizing that enables one to identify the lawful patterns behind many phenomena. In contrast unlawfulness might be thought to refer to information that might be ambiguous, since the outcome may vary, and as such is unpredictable. Hence identification of lawful patterns is more difficult for ambiguous information.

Thus, it might be this unlawfulness (ambiguity) of information to be processed, as put by Baron-Cohen (2009), which is posing extra difficulty for the ASD group for the repairman task in the current study. This explanation is more closely related to what might be predicted in relation to the disordered complex information processing theory. It is also possible that ambiguity is one way to define complexity; leading processing to be slowed during ambiguous tasks (Minshew et al., 2008). Our eye movement data are particularly informative regarding the timing of processing differences, and they correspond with one of our previous scene perception studies in which participants with ASD showed delayed orienting in a more ambiguous task, where participants had to decide "which one is weird" from a pair of almost identical scenes. We also observed an absence of such a delay using the same materials but with a simpler task that required participants to "spot the difference" between the two scenes (Au-Yeung et al., 2011; Benson et al., 2012).

In conclusion, we recognise that we did not set out to vary the different levels of ambiguity in our different perspective tasks. Moreover, a simpler and non-ambiguous control condition (e.g., "look for red squares") would have allowed us to examine in more detail whether the differences between the TD group and individuals with ASD are directly related to ambiguity. These are limitations that need to be fully addressed in future work. However, that aside, our ASD's were clearly able to take the perspective of another person, which would not have been predicted by the original ToM view. Importantly, it also appears that some perspectives may be more complex than others, and what makes that the case may be the ambiguity associated with the task. We would therefore like to conclude from the current, and our previous, findings that increasing ambiguity in the way information could be interpreted would be predicted to result in an early processing delay in ASD. We do not however think this delay would be restricted to empathizing tasks, but, and in line with the disordered complex information processing theory, we predict that such delays would be observed across a range of processing domains. Potentially, at an applied level, our findings suggest that inefficient processing of ambiguous information could contribute to some of the observed communication and social deficits in ASD.

Acknowledgments

References

Au-Yeung, S. K., Benson, V., Castelhano, M., & Rayner, K. (2011). Eye movement sequences during simple versus complex information processing of scenes in autism spectrum disorder. *Autism Research and Treatment*. doi:10.1155/2011/657383.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author. doi:10.1176/appi.books.978089042334

Baron-Cohen, S. (2001). Theory of mind in normal development and autism. *Prisme, 34*, 174-183.

Baron-Cohen, S. (2009). Autism: the empathizing-systemizing (E-S) theory. Annals of The New York Academy of Sciences, *1156*(1), 68-80.doi:10.1111/j.1749-6632.2009.04467.x

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism. *Journal of Child Psychology and Psychiatry*, *42*, 241-251. doi: 10.1111/1469-7610.00715

Baron-Cohen, S., Wheelwright, S., Robinson, J., & Woodbury-Smith, M. (2005). The Adult Asperger Assessment (AAA): *A diagnostic method. Journal of Autism and Developmental Disorders*, *35*, 807 - 819.

Benson, V., Castelhano, M., Au-Yeung, S. K., & Rayner, K. (2012). Eye movements reveal no immediate "WOW" ("whichone's weird") effect in autism spectrum disorder. *The Quarterly Journal of Experimental Psychology*. doi: 10.1080/17470218.2011.644305

Benson, V., Piper, J., & Fletcher-Watson, S. (2009). Abnormal saccadic sampling in autistic spectrum disorder. *Neuropsychologia*, *47*, 1178-1182. doi: 10.1016/j.neuropsychologia.2008.11.019

Henderson, J. M., Malcohm, G. L., & Schandl, C. (2009). Searching in the dark: Cognitive relevance drives attention in real-world scenes. *Psychonomic Bulletin& Review*, *16*, 850-865. doi: 10.3758/PBR.16.5.850

Kaakinen, J. K., & Hyönä, J. (2008). Perspective-driven text comprehension. *Applied Cognitive Psychology, 22*, 319-334. doi: 10.1002/acp.1412

Kaakinen, J. K., Hyönä, J., & Viljanen, M. (2011). Influence of a psychological perspective on scene viewing and memory for scenes. *Quarterly Journal of Experimental Psychology, 64*, 1372-1387.doi: 10.1080/17470218.2010.548872

Lord, C., Rutter, M., DiLavore, P., & Risi, S. (2001). *Autism diagnostic observation schedule (ADOS) manual*. Los Angeles, CA: Western Psychological Services.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview–Revised: A revised version of the diagnostic Interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*, 659 – 685.

Loth, E., Gomez, J. C., & Happé, F. (2011). Do high-functioning people with Autism Spectrum Disorder spontaneously use event knowledge to selectively attend to and remember context-relevant aspects in scenes*? Journal of Autism and Developmental Disorders*, *41*, 945-961. doi: 10.1007/s10803-010-1124-6

Minshew, N. J., & Goldstein, G. (1998). Autism as a disorder of complex information processing. *Mental Retardation and Developmental Disabilities Research Reviews, 4,* 129-136.doi: 10.1002/(SICI)1098-2779(1998)4:2<129::AID-MRDD10>3.0.CO;2-X

Minshew, N. J., Goldstein, G.& Siegel, D. J. (1997). Neuropsychologic functioning in autism: Profile of a complex information processing disorder. *Journal of the International Neuropsychological Society, 3*, 303-316.

Minshew, N. J., & Williams, D. L. (2008). Brain behavior connections in autism.*Speaker's Journal, 8*, 25 – 43.

Minshew, N. J., Williams, D. L., & McFadden, K. (2008). Information processing, neural connectivity, and neuronal organization. In A. W. Zimmerman (Ed.), *Autism: Current Clinical Neurology* (pp. 381-405). Totowa, NJ: Humana Press.

Pichert, J. W., & Anderson, R. C. (1977). Taking different perspectives on a story. *Journal of Educational Psychology*, *69*, 309 -315. doi: 10.1037/0022-0663.69.4.309

Rayner, K. (2009). The Thirty Fifth Sir Frederick Bartlett Lecture: Eye movements and attention during reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*, 1457-1506.

Senju, A. Southgate, V. White, S., & Frith, U. (2009). Mindblind eyes: an absence of spontaneous theory of mind in Asperger Syndrome. *Science*, *325*(5942), 883–885 .doi: 10.1126/science.1176170

The Psychological Corporation (1999). *Wechsler Abbreviated Scale of Intelligence (WASI) Manual*. San Antonio, TX: The Psychological Corporation.

World Health Organisation (1992). *International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10)*. Geneva: WHO.

Yarbus, A. L. (1967). Eye Movements and vision. New York: Plenum.

Figure 1. An example of a house scene. There are an equivalent number of interest items of each type. Each target was outlined by freehand with resulting regions of interest shown in yellow. The burglar targets were valuable items including 1) the laptop and 2) the printer. The repairman targets were features that need fixing including 1) the broken curtain rail and 2) the damaged radiator. It must also be noted that the valuable items were not restricted to technology items (a class of stimuli that individuals with ASD are known to favour) but also included non-technological items such as money, purses, handbags etc.

Figure 2. The magnitude of the relevance effect for total viewing time (top left), total number of fixations, (bottom left), elapsed time to target (top right), and fixation count to target (bottom right). Error bars represent *SEM*. Relevance effect is the difference in value between schema-relevant and irrelevant items for each of these dependent measures.