Memory for emotionally arousing events over time in autism spectrum disorder

Katie L Maras
Sebastian B Gaigg
Dermot M Bowler

Autism Research Group
City University London

In press with *Emotion*
Abstract

Emotionally arousing events are typically better remembered and more resistant to forgetting than neutral events. Findings from word list paradigms suggest that this may not hold for individuals with Autism Spectrum Disorder (ASD), who also tend to be less accurate as eyewitnesses under some circumstances. To test whether attenuated effects of arousal on memory may be responsible for poorer eyewitness testimonies in ASD, we asked adults with and without the disorder to view either arousing or neutral versions of a narrated slide sequence (Experiment 1) or video clip (Experiment 2) before assessing their memory for the material. Both groups exhibited increases in psychophysiological arousal during the arousing as compared to the neutral version of the narratives, and both groups also demonstrated a memory advantage for the arousing events. Contrary to predictions, these observations indicate that stimulus induced arousal modulates memory for naturalistic events relatively typically in ASD.

Key Words: Autism Spectrum Disorder, Emotion, Arousal, Eyewitness, Memory, Delay

Katie Maras, City University London. Katie.Maras.1@city.ac.uk
It is now well established that emotionally arousing events are better remembered and less likely forgotten than equivalent neutral events (e.g. Bornstein, Liebel, & Scarberry, 1998; Christianson & Loftus, 1991; Christianson, 1992; Heuer & Resiberg, 1990). Witnessed criminal events are often emotionally arousing, and witnesses can be asked by the police and other legal officials to recall what they have seen on a number of occasions, ranging from immediately after witnessing the event to days, months, or even years later. Maras and Bowler (2010) recently found that individuals with Autism Spectrum Disorder (ASD) were significantly less accurate in their eyewitness reports for a negative emotional event than comparison individuals when interviewed with a Cognitive Interview (Fisher & Geiselman, 1992). Here we ask whether the inaccuracies in eyewitness reports in ASD may be the result of abnormalities in how the emotional nature of criminal events modulates memory in this disorder.

Autism Spectrum Disorder is a set of pervasive developmental conditions that are clinically defined by abnormalities in reciprocal social and communicative behaviours and an inflexible adherence to routinised patterns of thought and behaviour (American Psychiatric Association, 2000). ASD affects approximately 1% of the population but a number of risk factors indicate that individuals with ASD may be over-represented within the Criminal Justice System either as witnesses or victims of crime (Hare, Gould, Mills & Wing, 1999; Petersilia, 2001). For example their diminished insight into what others are thinking can lead to exploitation by others (Howlin, 1997) and their repetitive and stereotyped interests can lead them to frequent places (e.g. train stations) where crimes are more common (Allen, Evans, Hider, Hawkins, Peckett & Morgan, 2008). This literature indicates an overwhelming need for research to examine eyewitness testimony in ASD, particularly because the very sparse work in this area to date suggests that individuals with ASD recall a previously witnessed event less completely and/or less accurately than comparison groups (Maras & Bowler, 2010, 2011; McCrory, Henry & Happé, 2007).

The defining reciprocal social impairments of ASD are *inter alia* characterised by difficulties with emotion related processes such as understanding, empathising and reciprocating emotional expressions in others (e.g. Dawson, Hill, Spencer & Galpert, 1990; Hobson, 2002; Kasari, Sigman, Mundy, & Yirmiya, 1990). It is therefore possible that people with ASD do not exhibit the same memory advantage for arousing as compared to neutral events as typical participants do, thus contributing to the poorer eyewitness testimonies in this disorder. To date, only four studies have examined whether individuals with ASD exhibit a typical memory advantage for emotionally significant information and the results from these are rather mixed. Beversdorf and colleagues (Beversdorf, Anderson, Manning, Anderson, Nordgren et al., 1998) found that high-functioning adults with ASD did not show enhanced memory for emotionally charged compared to neutral sentences like typically developed adults do. Similarly Deruelle, Hubert, Santos and Wicker (2008) reported no effect of emotional content on memory in an ASD group when positive, negative and neutral images were used as stimuli. By contrast, South and colleagues (South, Ozonoff, Suchy, Kesner, McMahon, et al 2008) found no differences between individuals with and without ASD in terms of their enhanced memory for emotionally salient as compared to neutral words. Finally, Gaigg and Bowler (2008) also failed to note differences between ASD and comparison individuals when assessing memory for emotionally charged and neutral words on an immediate test of memory. However, when these authors assessed memory again following a 1-hour and 1-day delay, the advantage for emotional material had faded for the ASD group whilst it had increased for the typical comparison group.

Together the evidence concerning memory for emotional material in ASD would seem to suggest that the emotional nature of witnessed events (e.g. accidents or crimes) may
not enhance the memory of witnesses with ASD as reliably as that of typical witnesses, particularly if memory for the event is probed following long delays. What may further exacerbate poor eyewitness testimonials in ASD is the possibility that witnessed events do not elicit the same kind of emotional responses in this disorder in the first place, thus altering not only how the event is encoded into memory but also what is attended to and hence encoded. Again, the evidence in relation to this issue is somewhat mixed. On the one hand, several studies indicate that individuals with ASD exhibit relatively typical psychophysiological responses, such as increases in Galvanic Skin Responses (GSR), to emotionally salient stimuli (e.g. Ben Shalom, Mostofsky, Hazlett, Goldberg, McLeod, et al., 2003; Gaigg & Bowler, 2008; Salmond, de Haan, Friston, Gadian, Vargha-Khadem et al., 2003), which would suggest that they may orient relatively typically toward emotionally salient events. On the other hand, there are also reports of differences in the emotional responses of individuals with ASD (Ben Shalom et al., 2003; Blair, 1999; Bölte, Feineis-Matthews & Poustka, 2008) particularly when witnessing others in distress (Corona, Dissanayake, Arbelle, Wellington & Sigman, 1998). Moreover, there is considerable behavioural evidence that emotionally salient information does not capture the attention of individuals with ASD typically (Ashwin, Wheelwright & Baron-Cohen, 2006; Corden, Chilvers & Skuse, 2008; Gaigg & Bowler, 2009; but see South et al., 2008) and that they spend less time attending to people’s faces when viewing complex social scenes (Klin, Jones, Schultz, Volkmar & Cohen, 2002), which is particularly relevant to eyewitness reports (see Maras & Bowler, 2010). Thus, differences in how individuals with ASD respond and attend to emotional information may compound, or even be responsible for atypicalities in how witnessed events are encoded and later retrieved from memory.

In order to address what role emotional factors might play in shaping eyewitness reports of individuals with ASD, we took advantage of a well established experimental paradigm developed by Heuer and Reisberg (1990), which involves presenting participants with a narrated slide-show that is either entirely neutral or includes an emotionally salient event (a car accident). Memory for the narrative is assessed following a delay of several days through free recall and forced-choice recognition tests and it is consistently found that memory for the emotional events in such paradigms is enhanced (see Reisberg & Heuer, 2004 for a review). To test the prediction that memory is atypically modulated by emotional factors in ASD, we replicated the method of Heuer & Reisberg (1990) in Experiment 1 and employed more ecologically valid video material in Experiment 2. Furthermore, to ascertain whether individuals with ASD exhibit atypical emotional responses to relevant events, GSR and Heart Rate was monitored in Experiment 1 and Experiment 2 respectively.

**Experiment 1**

**Method**

**Participants**

Twenty-seven individuals with a diagnosis of ASD (20 male; 7 female) and 29 typically developed comparison individuals (24 male; 5 female) were recruited for this experiment. Participants were randomly allocated to either the ‘Emotional’ or ‘Neutral’ story version of the experiment (see below), with the constraint that all sub-groups were matched in terms of chronological age and Wechsler Full-Scale IQ (all group and story version main effect and interaction $F$s < .45, $p$s > .77). Since 8 ASD and 10 comparison individuals failed
to return relevant memory measures following their initial participation, full data sets were only available for 19 participants in each group. Table 1 provides relevant psychometric data for this final sample of participants as a function of experimental condition.

Table 1

Summary of Age and Wechsler (1981) IQ characteristics of the ASD and Comparison Group as a function of experimental condition for Experiment 1 (there are no significant differences between groups or conditions)

<table>
<thead>
<tr>
<th>Condition</th>
<th>ASD (n = 19)</th>
<th>Comparison (n = 19)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Emotional version (n = 9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>35.2</td>
<td>13.2</td>
</tr>
<tr>
<td>Verbal IQ</td>
<td>109.7</td>
<td>18.0</td>
</tr>
<tr>
<td>Performance IQ</td>
<td>112.6</td>
<td>19.3</td>
</tr>
<tr>
<td>Full-scale IQ</td>
<td>111.3</td>
<td>19.2</td>
</tr>
<tr>
<td>Neutral version (n = 10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>37.4</td>
<td>15.6</td>
</tr>
<tr>
<td>Verbal IQ</td>
<td>107.8</td>
<td>18.1</td>
</tr>
<tr>
<td>Performance IQ</td>
<td>107.8</td>
<td>20.8</td>
</tr>
<tr>
<td>Full-scale IQ</td>
<td>107.9</td>
<td>20.5</td>
</tr>
</tbody>
</table>

ASD participants were diagnosed according to DSM-IV criteria (American Psychiatric Association, 2000) by experienced clinicians and for 15 of the 19 individuals an assessment with the Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, DiLavore & Risi, 1999) revealed a range of scores on the Communication (range = 1-5, mean = 2.38, SD = 1.12; cut off = 2) and Reciprocal Social Interaction domains (range = 3-12, mean = 7.08, SD = 3.55; cut off = 4) that were largely consistent with their clinical diagnosis. Three participants did not meet the research cut off for the combined ADOS score of 7 (one participant because they failed to reach the Communication cut off, and two because they failed to reach the reciprocal social interaction cut off). Since clinical records were available to confirm their diagnosis we retained these participants in all analyses. Similarly the four ASD participants for whom no ADOS observations were available had a clear clinical statement of their diagnosis and were therefore also retained in the analysis. Comparison
participants were all in good health and reported no family history of psychiatric or neurological illness.

Materials & Design

Following Heuer and Reisberg (1990), we constructed two versions of a 12-image, narrated slideshow. The first three slides were identical in both versions and depicted a mother and son leaving home (slide 1), crossing a street (slide 2) and walking along a path (slide 3) whilst the accompanying narrative told the listener that the protagonists set off in the morning to visit the son’s father at work. The next 5 slides differed between the two versions. In the ‘Neutral’ version, the father was identified as head mechanic in a local garage (slide 4) and the story explained that a car had been towed in earlier to be fixed (slides 5 & 6) and that the father had been working on a racing car engine for several hours in the back room (slides 7 & 8). In the ‘Emotional’ version, the father was identified as chief surgeon in a local hospital (slide 4) and participants were informed that a car-crash victim was undergoing emergency surgery (slides 5 & 6) whilst the father was monitoring the critical condition of a young infant who had been injured during the crash (slides 7 & 8). Slides 9-12 were once again identical in the two versions and showed the mother leaving the father’s work place (slide 9) to phone her boss from a phone-booth (slides 10 & 11) before hailing a cab to make her way home (slide 12). Appendix 2 provides a summary description of all experimental materials.

Slides were projected at a rate of 10 seconds per slide with 2 second blank intervals between successive slides. The accompanying narrative was pre-recorded by a female speaker and presented at slide onset. Memory for the slideshow was assessed following a delay of several days through both a free recall and a forced choice recognition test, which were mailed to the participants. Instructions enclosed with the mailing asked participants to complete two memory tasks related to the slideshow and that they should begin by writing down everything they could remember about it in as much detail as possible. They were then asked to complete a 4 alternative forced choice test comprising 10 questions for each of the 12 slides in the story (e.g. ‘Who are the surgeons trying to save?’, ‘What was wrong with the car?’). They were also asked to rate their confidence in the answers they chose on a scale from 1 (guessing) to 5 (completely sure) by writing relevant numbers next to their choices. Finally, participants were told that it was important to answer the questions in sequence and to not back-track to change any answers.

Throughout the presentation of the slideshow, Galvanic Skin Responses (GSRs) were monitored via surface electrodes attached to the palmar surfaces of the medial phalanges of the first and third digits of the participants’ non-dominant hand. A PowerLab system (ADInstruments Ltd. 2004a) running Chart 5 software (ADInstruments Ltd. 2004b) acquired GSR data at a rate of 1 kHz, and E-prime software interfaced with the PowerLab unit in order to demarcate the GSR trace with relevant slide onsets. The maximum positive deflections in signal amplitudes were computed for each of the 12 slides by subtracting the maximum value

---

1 We are grateful to the authors for providing us with their original materials. Slides 1-3 and 9 – 12 in the current materials are adapted from the original Heuer & Reisberg (1990) publication but slides 4 – 8 of each version of the slideshow were sourced from various websites for the purposes of the present study.
within 8 seconds following slide onset from the value at slide onset. A square-root transformation was applied to normalise the data and visual inspection of all raw data confirmed the absence of movement artefacts.

**Procedure**

Participants were tested individually and were informed that they would be asked to watch a short narrated slideshow about a boy visiting his father at work. All were warned that the slideshow may include emotionally disturbing images such as pictures of somebody getting hurt. The instructions explained that two electrodes attached to the hand would measure their sweat responses throughout the slideshow and that a baseball cap equipped with cameras would monitor eye movements. Once informed consent was obtained the room was darkened and the slideshow started. Immediately after the slideshow, participants were debriefed and asked to rate the narrative in terms of how emotionally stimulating the story was (on a 1-10 scale; 1 = not at all emotionally stimulating, 10 = extremely emotionally stimulating). No mention was made at any point that memory for the slideshow would be assessed later.

**Coding and Preliminary Analysis**

In order to quantify participants’ free recall attempts, we followed Heuer and Reisberg’s (1990) procedure of segmenting participant’s verbal reports into individual phrases and then classifying these as either reflecting information from the beginning (Slides 1-4), middle (Slides 5-7) or end (Slides 8-12) of the slideshow. To classify phrases systematically, 39 discrete details were identified that participants could recall from the beginning of the narrative, 26 details that they could recall from the critical middle segment and 39 details that they could recall from the end. These details were identified on the basis of the items probed on the forced-choice recognition questionnaire and a rater blind to participant diagnosis coded all free recall phrases as either correct or incorrect depending on whether the information accurately reflected one of the identified details or not. A second independent rater scored eight randomly selected interview transcripts (two in each group x condition) against the transcription of the slides and the resulting Person’s correlations of the two coders’ scores were: $r_{\text{correct}} = .92, p < .001$, $r_{\text{incorrect}} = 0.74, p < .05$.

**Results and discussion**

**Memory Data**

Prior to analysing participants’ recall and forced choice recognition performance, it was important to ascertain whether these behavioural measures might be confounded by

---

2 For pilot purposes we recorded eye gaze throughout the experiment.
systematic differences in the time it took participants to complete the relevant memory tasks. Based on the dates participants provided on the forms they returned, ASD individuals in the neutral condition completed the memory tasks on average 13 (range 4 – 33) days after seeing the slideshow whilst those in the emotional condition took on average 12 (range 4 – 31) days. The respective values for the comparison group were 13 (range 4 – 29) and 9 (range 4 – 22). None of these differences was significant (all $F$s $< 0.9$), nor was there any correlation between the time taken to complete the tasks and any of the memory measures reported below (all $r$s $< 0.1$).

Table 2 summarises the free recall and forced choice recognition data as a function of group and shows that the emotional version of the story yielded higher rates of recall in both participant groups, particularly with respect to the critical middle segment of the slideshow. Analysis of the recall data yielded main effects of Story Version, $F (1, 34) = 4.28$, $p < .05$, $r = .33$, and Segment, $F (2, 68) = 3.32$, $p < .05$, $r = .22$, as well as a two-way interaction between these factors, $F (2, 68) = 10.19$, $p < .001$, $r = .36$. Neither the main effect of group, $F (1, 34) = 3.04$, $ns$, $r = .29$, nor any interactions involving group (all $F$s $< 1.2$, $ps > .32$) were significant. As Table 2 suggests, the two-way interaction resulted from significantly more information being recalled in the emotional condition than in the neutral condition from the critical middle segment of the story, $t (36) = 3.36$, $p < .01$, Cohen’s $d = 1.07$, but not from the beginning, $t (36) = .33$, $ns$, Cohen’s $d = .11$, or end segments, $t (36) = 1.55$, $ns$, Cohen’s $d = .49$.

Table 2

Free recall measure: The average number of phrases reflecting information from either the Beginning, Middle or End of the emotional or neutral slideshows as a function of Group for Experiment 1

<table>
<thead>
<tr>
<th>Slideshow</th>
<th>Free Recall</th>
<th>Forced Choice Recognition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Number of Phrases)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ASD ($n = 19$)</td>
<td>Comparison ($n = 19$)</td>
</tr>
<tr>
<td>Emotional</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beginning</td>
<td>4.22</td>
<td>3.19</td>
</tr>
<tr>
<td>Middle</td>
<td>5.89</td>
<td>5.21</td>
</tr>
<tr>
<td>End</td>
<td>4.11</td>
<td>3.69</td>
</tr>
<tr>
<td>Neutral</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Analysis of participants’ forced choice recognition performance mirrored the results of the free recall measure revealing a significant interaction between Story Version and Segment, $F(2, 66) = 3.24, p < .05, r = .22$, that was due to participants in the emotional slideshow condition performing better than those in the neutral condition particularly on the critical middle segment of the story, $t(36) = 3.25, p < .05$, Cohen’s $d = 1.04$. Performance on questions concerning the beginning, $t(36) = 1.48, ns$, Cohen’s $d = 0.49$, and end, $t(36) = 0.93, ns$, Cohen’s $d = 0.30$, of the story once again did not differ in the two experimental conditions, and a lack of main effects or interactions involving the group factor (all $F$s $< = 2.1, p$s $> .13$) suggests that the patterning of performance was similar across groups.

Next we analysed errors, and findings remained largely the same. There was a significant Story Version x Segment interaction, $F(2, 68) = 3.13, p < .05, r = .21$, which was again qualified by a memory advantage (with fewer errors) for the critical middle segment of the arousing version, $t(19) = 2.30, p < .05$, Cohen’s $d = .77$, and no difference in errors reported between the beginning, $t(33) = 1.30, ns$, Cohen’s $d = .42$, or end segments, $t(36) = 1.35, ns$, Cohen’s $d = .43$. There was also a main effect of Segment, $F(2, 68) = 8.35, p < .001, r = .33$, which was due to a significant decrease in errors between the beginning and middle segments, $t(37) = 4.21, p < .001$, Cohen’s $d = 1.03$, and a significant increase in reporting of errors between middle and end segments, $t(37) = 2.06, p < .05$, Cohen’s $d = .47$. There was a marginal main effect of Group, $F(1, 34) = 3.53, p = .07, r = .31$, with the ASD participants reporting somewhat more errors (mean = 3.79 SD = 3.07) than the comparison group (mean = 2.37, SD = 1.38). Follow-up between subject $t$-tests showed that the ASD group made significantly more errors than the comparison group only for the middle segment of the emotional version of the story, $t(16) = 2.13, p < .05$, Cohen’s $d = 1.00$. No other main effects or interactions for errors were significant (all $F$s $< 1.07, p$s $> .31$).

**Psychophysiological Responses and Subjective Ratings**

In order to determine whether the emotional version of the slideshow was indeed experienced as more emotional than the neutral version, a 2 (Group) x 2 (Story Version: Emotional vs. Neutral) univariate analysis of variance was carried out on the participants’ subjective ratings of the slideshows. Results confirmed a main effect of Story Version, $F(1, 34) = 33.23, p < .001, r = .70$, with participants in the emotional condition rating the story as more emotional (ASD mean = 4.78, SD = 2.44; comparison mean = 4.89, SD = 1.45) than participants in the neutral condition (ASD mean = 2.20, SD = 1.62; comparison mean = 1.30, SD = .67). Neither the main effect of Group $F(1, 34) = .54, ns, r = .13$, nor the interaction between the factors, $F(1, 34) = .89, ns, r = .16$, were significant, suggesting that the

<table>
<thead>
<tr>
<th>Segment</th>
<th>Story Version 1</th>
<th>Story Version 2</th>
<th>Story Version 3</th>
<th>Story Version 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginning</td>
<td>3.00</td>
<td>2.49</td>
<td>3.10</td>
<td>2.60</td>
</tr>
<tr>
<td>Middle</td>
<td>1.10</td>
<td>1.10</td>
<td>0.50</td>
<td>1.08</td>
</tr>
<tr>
<td>End</td>
<td>2.70</td>
<td>2.83</td>
<td>3.00</td>
<td>2.58</td>
</tr>
</tbody>
</table>

*a Significant difference between story versions for recall, $p < .01$ and recognition, $p < .05$. 
experimental manipulation was similarly effective for both groups. An analysis of the GSR data further corroborated this conclusion.

Average GSR amplitudes elicited by the beginning (slides 1-4), middle (slides 5-7) and end (slides 8-12) segments of the two versions of the narrative were calculated. An initial inspection of the means totalled across all segments revealed that both groups demonstrated higher GSR amplitudes for the emotional version (mean ASD = .36, SD = .20; mean comparison = .27, SD = .21) compared to the neutral version (mean ASD = .32, SD = .18; mean comparison = .24, SD = .18). Although these data are characterised by considerable variability both within and between groups, analysis of the data confirmed a two-way interaction between Story Version and Segment, \( F(2, 66) = 3.24, p < .05, r = .22 \), and no other main effects or interactions (all \( F_s < 2.1, p_s > .13 \)). As expected, the interaction was due to significant differences in GSR amplitudes across the three segments of the emotional, \( F(2, 32) = 5.73, p < .01, r = .39 \) (beginning mean = .34, SD = .23; middle mean = .38, SD = .28; end mean = .23, SD = .18), but not the neutral version of the slideshow, \( F(2, 34) = .11, ns, r = .06 \) (beginning mean = .28, SD = .22; middle mean = .27, SD = .19; end mean = .29, SD = .25). Given the variability of these data, it is useful to note that an analysis of the entire sample of 27 ASD and 29 typical comparison individuals who had originally participated in the experiment yielded virtually identical results with a significant two-way interaction between Story Version and Segment, \( F(2, 102) = 3.92, p < .05, r = .19 \), and no other main effects or interactions (all \( F_s < 2.2, p_s > .14 \)).

The findings from Experiment 1 indicate that the ASD group’s recall was similarly enhanced by the emotional nature of the middle segment of the arousing version of the narrated slideshow as the comparison group’s was. Both groups recalled significantly more correct details from the emotional slides compared to the neutral slides and both groups exhibited similar psychophysiological and subjective emotional responses to the emotional version of the narrative. Because these findings are rather unexpected and contrast the observations of several previous studies (e.g. Beversdorf et al., 1998; Deruelle et al., 2008; Gaigg & Bowler, 2008), we sought to substantiate them with a larger sample and more ecologically valid materials in Experiment 2. Thus, Experiment 2 employed a video clip rather than static slides as the to-be-remembered event and in order to examine whether the patterning of the modulation of memory by arousal differentiated the ASD and comparison groups over time, recall was also tested at three time points.

**Experiment 2**

**Method**

*Participants*

Twenty-four typically developed and twenty-four ASD participants were recruited for Experiment 2 in a similar manner to that described for Experiment 1. Their details are summarised in Table 3. There were no differences between groups or arousal conditions, or group x arousal interactions for the measures of IQ or age (all \( F_s < .59, p_s > .68 \)). Four

---

3 GSR data for 1 ASD participant in the emotional condition were not available due to a technical fault.
comparison participants and seven ASD participants had also taken part in Experiment 1. Scores on the ADOS Communication (range = 1-5, mean = 2.71, SD = 1.45; cut off = 2) and Reciprocal Social Interaction subscales (3-12, mean = 7.23, SD = 2.79; cut off = 4) were available for 22 of the 24 ASD participants and again these were largely in line with the independent clinical diagnosis from health professionals in the UK. Two participants did not meet the research cut off for the combined ADOS score of 7 (one participant because they failed to reach the Communication cut off, and one because they failed to reach the reciprocal social interaction cut off). Again, since clinical records were available to confirm their diagnosis we retained these participants in all analyses. Similarly the two ASD participants for whom no ADOS observations were available had a clear clinical statement of their diagnosis and were therefore also retained in the analysis.

Table 3

Age and IQ scores for the ASD and comparison groups, within each arousal condition (standard deviations in parentheses) for Experiment 2 (there are no significant differences between groups or conditions)

<table>
<thead>
<tr>
<th>Condition (Clip version)</th>
<th>ASD (N = 24)</th>
<th>Comparison (N= 24)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arousing (N = 24)</td>
<td>(n = 12)</td>
<td>(n = 12)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>40.00 (11.98)</td>
<td>43.33 (10.40)</td>
</tr>
<tr>
<td>Verbal IQ</td>
<td>111.67 (13.77)</td>
<td>109.92 (14.90)</td>
</tr>
<tr>
<td>Performance IQ</td>
<td>105.75 (14.83)</td>
<td>109.00 (15.27)</td>
</tr>
<tr>
<td>Full-scale IQ</td>
<td>109.92 (15.14)</td>
<td>110.33 (15.92)</td>
</tr>
<tr>
<td>Neutral (N = 24)</td>
<td>(n = 12)</td>
<td>(n = 12)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>41.00 (12.55)</td>
<td>43.25 (14.40)</td>
</tr>
<tr>
<td>Verbal IQ</td>
<td>114.18 (9.42)</td>
<td>111.33 (14.47)</td>
</tr>
<tr>
<td>Performance IQ</td>
<td>109.91 (15.28)</td>
<td>106.42 (10.64)</td>
</tr>
<tr>
<td>Full-scale IQ</td>
<td>113.64 (11.27)</td>
<td>109.83 (13.66)</td>
</tr>
</tbody>
</table>

Materials

Participants viewed a short scene from a certificate 15 rated film (UK accreditation) that had already been successfully used for a similar purpose in previous work (Bornstein et al, 1998). Both versions were set in a graveyard and began by showing a kneeling man laying flowers on a grave and a priest walking around the graveyard. The middle segment then began and, as in Experiment 1, differentiated the emotional and neutral versions of the film. In the emotional version a male protagonist approached, pulled a gun from his cloak and shot the man who was kneeling down in the back of the head. He then aimed the gun at the priest who was watching before backing slowly away. The neutral version contained the same characters in the same setting but the plot concentrated on non-violent events that focused on the priest, who in this version was going about his business without witnessing the crime.
Both versions then concluded with identical end segments showing the priest going over to the injured man (no injury was visible but implied) and reading him the last rites (See Appendix B for further details). Thus, although both versions of the film can be regarded as negative in the sense that they were about a man being fatally injured, only the arousing version illustrated this explicitly and graphically. Both versions lasted a total of 91 seconds. Permission was obtained from the film company to edit and use the film for the purpose of this study.

Throughout the video clip, participants’ heart rate was measured using conductive adhesive electrodes attached to each wrist with a reference electrode attached to the elbow. Electrocardiograms (ECG) were monitored using the same PowerLab system (ADInstruments Ltd. 2004a) as in Experiment 1 and Chart 5 software (ADInstruments Ltd. 2004b) was used to compute beats-per-minute (bpm) for each of the three segments of the clip (beginning, middle and end). The use of heart rate rather than GSR as the physiological measure in this experiment was more appropriate because of the continuous nature of the video-clip.

Procedure

Participants were tested individually and informed that the purpose of the study was to assess their patterns of physiological reactions to neutral and arousing video stimuli, and how their physiological responses related to their subjective experiences of arousal. They were informed that they were about to watch a short video and that they should not specifically try to remember the details but instead simply watch it as if they were watching an event in real life. The ECG electrodes were attached and participants were asked to relax into a comfortable position in front of the computer screen and the video clip was then played. Immediately after the clip had finished playing and the heart rate equipment was removed, participants were asked to rate how arousing they found the clip on a scale of 1 to 5 (1 = not at all arousing; 5 = very arousing). Next participants were given a surprise memory test. They were asked to write down everything they could recall from the clip in as much detail as they could, with no time constraints on how long they had to do this. Following some unrelated tasks lasting around an hour, participants were again asked to write down everything they could remember from the clip, including all the details they had previously written and any extra details they might remember. Participants were then given a sealed envelope containing instructions for a third recall attempt and a self addressed envelope for returning this by post. They were asked not to open or look at the contents until the following day. Receipt of completed forms and follow-up phone calls confirmed that all participants had completed their recall forms the day after initial testing.

Coding and preliminary analyses

Similar to Experiment 1 a standardised coding procedure was used for quantifying participants’ free recall of the video. More specifically, the same discrete details reliably identified by Bornstein et al (1998) in each version of the film were used to code participants’ responses. There were 45 details in the arousing version and 43 details in the neutral condition. Both versions contained 18 details in the beginning segment and 14 details in the end segment. The arousing version of the middle segment contained 13 details and the neutral version contained 11 details (Appendix A). To enable comparisons between clip version
segments we calculated proportions separately for each participant by dividing the total number of reported details by the total possible number of details in that segment. If a participant reported a detail that was not seen in the clip at all, or was inaccurate this was coded as an error. A second independent rater scored eight randomly selected interview transcripts (two in each group x condition) against the video clip transcription and the resulting Person’s correlations of the two coders’ scores were: \( r_{\text{correct}} = .98, p < .0001 \), \( r_{\text{incorrect}} = 0.86, p < .01 \).

Tests for normality and outliers revealed one ASD participant who recalled a high proportion of details. Inclusion of this participant changed the findings for only one analysis, which is reported below. We analysed the data with the main question in mind of whether each group’s recall was differentially affected by arousal over the different delay periods.

**Results and discussion**

**Memory Data**

The free recall data of Experiment 2 are set out in Table 4. A 2 (group: ASD vs. comparison) x 2 (arousal: neutral vs. arousing clip versions) x 3 (segment: beginning vs. middle vs. end) x 3 (delay: immediate vs. one hour vs. one day) ANOVA was carried out to examine the data. As in Experiment 1, there was a main effect of arousal, \( F(1, 44) = 17.78, p < .001, r = .54 \), and significant segment x arousal interaction, \( F(2, 74) = 3.66, p < .05, r = .22 \), reflecting higher recall of details in the arousing (mean = .34, SD = .13) than in the neutral version (mean = .21, SD = .10), particularly in relation to the middle segment of the clip (all ts at all three time delays > 2.74, ps < .05). The recall data also yielded a marginally significant effect of group with ASD individuals recalling a smaller proportion of correct details (mean = .25, SD = .15) than the comparison group (mean = .31, SD = .11), \( F(1, 44) = 3.51, p = .068, r = .27 \). When the participant earlier identified as an outlier was removed from the data, this difference was significant, \( F(1, 43) = 6.62, p < .05, r = .37 \) (mean ASD = .23, SD = .12; mean comparison = .31, SD = .11). Importantly, however, there was no group x arousal interaction, \( F(1, 44) = 1.86, ns, r = .20 \). Neither the group x segment, \( F(2, 74) = 1.25, ns, r = .13 \), or group x segment x arousal interaction, \( F(2, 74) = 1.49, ns, r = .14 \) were significant, suggesting that the emotional salience of the video clip modulated memory relatively similarly in both groups.

In relation to the effect of time delay on recall, whilst there was not a main effect of delay, \( F(2, 74) = 1.97, p = .15, r = .16 \), there was a significant delay x arousal interaction, \( F(2, 73) = 3.66, p < .05, r = .22 \), whereby there was no difference in the proportion of recall over time for participants who viewed the arousing version (all ts < 1.00, ns), whereas participants in the neutral condition recalled significantly fewer details after one day (mean = .19, SD = .10) than they did immediately after watching the clip (mean = .23, SD = .10), \( t(23) = 2.76, p < .05 \), Cohen’s \( d = .35 \). The data were also characterised by a marginal delay x group interaction, \( F(2, 74) = 2.91, p = .06, r = .19 \), but no higher-order interactions, again suggesting that the emotional salience of the video clips modulated memory relatively similarly in both groups. Follow up t-tests revealed that whilst there was no difference between groups in the proportion of details recalled on the immediate or one day tests (ts < 1.41, ps > .17), the ASD group recalled marginally significantly fewer details than the
comparison group after one hour, $t (46) = 1.90$, $p = .06$, Cohen’s $d = .55$. These data are summarised in Table 4.

Table 4

Summary t-test results comparing the proportion of details recalled from the arousing and neutral versions of the video clip for ASD and comparison participants for each of the three time delays of recall for Experiment 2 (standard deviations are in parentheses)

<table>
<thead>
<tr>
<th>Video</th>
<th>ASD ($n = 24$)</th>
<th>Comparison ($n = 24$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Immediate</td>
<td>1 Hour</td>
</tr>
<tr>
<td>Arousing</td>
<td>.35 (.15)</td>
<td>.32 (.16)</td>
</tr>
<tr>
<td>Neutral</td>
<td>.17 (.08)</td>
<td>.16 (.09)</td>
</tr>
<tr>
<td>$t$</td>
<td>3.63</td>
<td>2.99</td>
</tr>
<tr>
<td>df</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>$p$</td>
<td>$&lt; .005$</td>
<td>$&lt; .01$</td>
</tr>
<tr>
<td>Cohen’s $d$</td>
<td>1.50</td>
<td>.91</td>
</tr>
</tbody>
</table>

We also analysed errors. A 2 (group) x 2 (arousal) x 3 (segment) x 3 (delay) ANOVA revealed a significant delay by group interaction, $F (1, 176) = 3.11$, $p < .05$, $r = .13$. Post-hoc tests revealed that this was due to the ASD group making significantly more errors on the immediate test, $t (30) = 2.56$, $p < .05$, Cohen’s $d = .74$, which substantiates the trend of increased errors in ASD participants observed in Experiment 1. There was no difference in error rates between groups, however, on the tests one hour, $t (46) = .15$, $ns$, Cohen’s $d = .04$, or one day later, $t (46) = .82$, $ns$, Cohen’s $d = .06$ suggesting that incorrectly recalled information, like correctly recalled details, are forgotten more rapidly by individuals with ASD. No other main effects or interactions for errors were significant (all $F$s < 2.69, $ps > .09$).

Psychophysiological Responses and Subjective Ratings
Analysis of participants’ subjective ratings of arousal revealed a significant main effect of arousal, $F(1, 176) = 4.97, p < .05, r = .32$: the arousing version yielded higher ratings (Mean = 2.85, SD = 1.14) than the neutral version (Mean = 2.17, SD = .96). There was no main effect of group, $F(1, 44) = .370, ns, r = 0.09$, nor a group x arousal interaction, $F(1, 44) = 0.37, ns, r = .10$.

Next we examined participant’s averaged heart rate for each version of the clip in the beginning, middle, and end segments. There was a significant main effect of segment, $F(2, 86) = 4.02, p < .05, r = .21$, but no segment x group interaction, $F(2, 86) = 1.93, ns, r = .15$, segment x arousal interaction, $F(2, 86) = .33, ns, r = .06$, or segment x group x arousal interaction, $F(2, 86) = .38, ns, r = .07$. Post-hoc paired $t$-tests to follow up the main effect of segment revealed that heart rate significantly dropped from the beginning (mean = 73.05, SD = 12.66) to the middle (mean = 72.24, SD = 11.94) segment, $t(46) = 2.47, p < .05, Cohen’s d = .07$, and significantly rose from the middle to the end (mean = 72.77, SD = 12.10) segment, $t(46) = 2.17, p < .05, Cohen’s d = .04$. There was no difference in heart rate between the beginning and end segments, $t(46) = 1.01, ns$, Cohen’s $d = .02$, indicating that the main effect of segment was due to a significant drop in the middle segment of the clip in both arousing and neutral conditions. Because we predicted that arousal would be modulated more in the arousing version of the clip for both groups we carried out planned comparisons across groups within each arousal condition. Within the arousing condition a similar pattern emerged: participant’s heart rate significantly decreased between the beginning (mean = 72.95, SD = 11.27) and middle (mean = 71.89, SD = 11.07) segments, $t(22) = 2.41, p < .05$, Cohen’s $d = .09$, and marginally significantly increased from the middle to end (mean = 72.55, SD = 10.85) segments, $t(22) = 1.93, p = .067, Cohen’s d = .06$, with no difference between the beginning and end segments, $t(22) = .99, ns$, Cohen’s $d = .04$. Within the neutral condition however, there was no difference in heart rate between any of the clip segments (beginning mean = 73.16, SD = 14.12; middle mean = 72.59, SD = 12.95; end mean = 72.98, SD = 13.43, all $t$s < 1.17, $p$s > .26).

A decrease in heart rate is thought to be part of an orienting response, and is a common reaction to unpleasant stimulation (e.g. Hare, Wood, Britain & Shadman, 1971). Increased cardiac activity is thought to be part of defensive response, evoked by strongly unpleasant or traumatic stimulus events (e.g. Christianson, 1987). Given that the stimuli used here were likely to be less overtly threatening than a real life situation it is unsurprising that they evoked an orienting response in our participants, and these findings are in-line with Heuer and Reisberg (1990), who found a downward turn in heart rate for the arousal group on the first slide from which the emotionally arousing events began. However what is of most interest in the present findings is that the ASD group exhibited similar orienting responses as the comparison group to the arousing clip version.

The reduced recall of correct details and increased reporting of errors on the immediate test by the ASD group is in line with previous work showing diminished free recall in this population (e.g. Bennetto, Pennington & Rogers, 1996; Bowler, Gardiner, Grice & Saavalainen, 2000; Maras & Bowler, 2011). In line with Experiment 1, the recall of both participant groups was equally enhanced in the arousing compared to the neutral version following a delay of one hour and one day for both groups.
General Discussion

The present research aimed to extend previous work on memory for emotional material in ASD (e.g. Gaigg & Bowler, 2008) in order to determine whether atypicalities in this domain may be responsible for relatively poorer eyewitness reports in this population (Maras & Bowler, 2010). Several lines of evidence led us to predict that when presented with eyewitness stimuli, the typical comparison group would remember emotional versions of events better and forget them less over time than neutral versions, whilst we expected no such modulation for a group of ASD individuals. This prediction was primarily based on three sets of findings. First, several studies indicate that individuals with ASD do not exhibit a typical memory advantage for emotionally salient material (Beversdorf, 1996; Deruelle et al. 2008), particularly when memory is assessed over time (Gaigg & Bowler, 2008). Second, there is evidence to suggest that individuals with ASD may not attend to the type of information (e.g. people in distress) that is critical in the context of eyewitness situations (e.g. Corona et al., 1998). And finally Maras and Bowler (2010) have shown that the Cognitive Interview, which encourages the mental reinstatement of contextual information such as the feeling one experienced at the time of witnessing an event, proves detrimental to the eyewitness testimonials of individuals with ASD.

Despite the converging evidence for our predictions, the observations from two experiments showed that both individuals with and without a diagnosis of ASD demonstrate enhanced memory for, and diminished forgetting rates of, emotionally salient as compared to neutral events. This contrasts the observations of previous studies on memory for emotional sentences (Beversdorf et al, 1998) pictures (Deruelle et al., 2008) and words (Gaigg & Bowler, 2008) in ASD, but is in line with another report on memory for emotional words in ASD by South and colleagues (2008). Thus, findings from studies which have examined memory for emotional material in ASD are rather inconsistent and at present it remains unclear what factors determine whether or not emotional factors modulate memory typically in ASD or not. The type of material, mode of presentation and delay between study and test varies considerably across the few relevant studies to date, and future studies should seek to vary these factors systematically. In addition, more work is needed at the more basic level of understanding the subjective and physiological components of emotional responses of individuals with ASD. In the present experiment ASD participants exhibited similarly enhanced levels of physiological responses (GSR and heart rate) to arousing as compared to neutral material as comparison participants, but again this is consistent only with some previous findings (e.g. Gaigg & Bowler, 2008; Ben Shalom et al, 2003; Salmond et al, 2003), but not others (e.g. Ben Shalom et al., 2003; Blair, 1999; Bölte, et al., 2008; Corona et al., 1998). Thus, more systematic work such as that by Bölte and colleagues (2008) is needed to resolve how individuals with ASD respond to, and subjectively experience emotional material. One avenue for future work that might prove particularly fruitful would be to examine orienting versus defensive responses to a variety of stimuli in ASD and to assess how these distinct responses modulate remembering in this group. Similar to the present study, Sigman, Dissanayake, Corona and Espinosa (2003) found that ASD children, like their matched typical counterparts, demonstrated an orienting response with decreased heart rate to affective social stimuli that were presented in a video clip. A study by Corona et al (1998), however, reported that their ASD children did not demonstrate an orienting response to seeing an experimenter appear in distress in real life. As Sigman et al (2003) point out, it would be worth considering the effect of watching a video versus seeing the events in real life.
on physiological responses (and their subsequent modulation with memory) by individuals with ASD.

Issues concerning emotions aside, both experiments provided evidence of relatively poorer recall abilities in ASD overall. In both experiments, the free recall reports of individuals with ASD were characterised by increased rates of errors and in Experiment 2 participants with ASD furthermore exhibited elevated rates of forgetting of correct details over time. A substantial body of empirical work shows that ASD is characterised by a profile of specific memory strengths and weaknesses. For example, whilst free-recall is often found to be diminished, when test procedures that provide more support for the studied material, such as cued recall or recognition, are used individuals with ASD usually show intact performance (see Bowler & Gaigg, 2008). Our findings are consistent both with this existing empirical work showing that free recall procedures pose particular difficulties for individuals with ASD (see Bowler & Gaigg, 2008), and with more applied previous research, demonstrating poorer eyewitness testimonials in ASD (Maras & Bowler, 2010, 2011). The preserved performance on the forced choice recognition test in Experiment 1 lends further support to the conclusion that supported test procedures yield relatively undiminished performance in ASD. Taken together these findings indicate that more generic recall difficulties, rather than specific abnormalities with recalling emotional information, may be responsible for the difficulties in recalling eyewitness events experienced by individuals with ASD. In this context, however, it is important to note that generalisation of the current findings to real-life eyewitness events is limited by a number of considerations. First, participants recalled details about a slide sequence or videoed event and were aware from the outset that the depicted events were staged, which could have attenuated arousal, or at least the overt social relevance of the event. Second, one needs to take into account the difference between laboratory research, where emotional arousal is assumed to reach a certain criterion, and real life situations, where emotional arousal might well exceed that criterion and consequently begin to have a negative effect on later memory (see Christianson, 1992). Third, we did not specifically match the neutral and arousing versions for distinctiveness so it could be argued that the arousing version was better recalled because it was more distinctive (but see Gaigg & Bowler, 2008). However, most criminal events are both arousing and distinctive, so for the purposes of applying this to real-life eyewitness testimony as well as for theoretical purposes, we chose to have emotional and neutral versions regardless of distinctiveness. Finally, we acknowledge the relatively small sample sizes used, particularly in Experiment 1. Whilst small sample size is usually noted in the context of limited power, this does not apply here given that we found the predicted effect. Nevertheless, it could be argued that in small datasets it is possible that one or two spurious individuals can make a big difference, however a close inspection of the data for our samples indicated that this was not a concern.

Despite these caveats, our findings make an important contribution both to the ASD and to the eyewitness literature. They show that arousal can modulate remembering in ASD for the kinds of stimuli that witnesses are likely to experience, which suggests that as eyewitnesses, individuals with ASD are as likely to demonstrate enhanced recall of details for emotionally arousing events as are typical witnesses.
References


ADInstruments. (2004b). Chart 5 v5.2 07. ADInstruments Pty Ltd.


Appendix A

Summary of the slideshow images together with transcripts of the accompanying narrative for Experiment 1

<table>
<thead>
<tr>
<th>Emotional Version</th>
<th>Neutral Version</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Image</strong></td>
<td><strong>Narrative</strong></td>
</tr>
<tr>
<td>Mother and son exiting home</td>
<td>This is Victoria and her son Phillip. They live on 85 Compton Close and are leaving the house early in this morning this Monday.</td>
</tr>
<tr>
<td>Mother and son waiting to cross road</td>
<td>They make sure that crossing Park Road is safe. Lately, cars have been driving by too fast on this street.</td>
</tr>
<tr>
<td>Mother and son walking along a path</td>
<td>Phillip is visiting his dad, Martin, at work today. It is a warm and sunny day for the end of October.</td>
</tr>
<tr>
<td>Outside view of ‘Mercury Motors’ garage</td>
<td>Martin is head mechanic at the nearby garage. He has been working at Mercury Motors for more than 8 years.</td>
</tr>
<tr>
<td>Front view of a parked car</td>
<td>Earlier in the day, this car had to be towed in. The owner said that he heard a grinding noise before the motor stopped.</td>
</tr>
<tr>
<td>Two mechanics working on car engine</td>
<td>The garage crew is trying to locate the problem. They discover that a broken screw has punctured the hydraulic tubing and are able to stop the oil leak.</td>
</tr>
</tbody>
</table>
Appendix B

Sequence of events in arousing and neutral clip versions for Experiment 2 (Bornstein et al., 1998)

Beginning segment - both versions (18 details)

1. Occurred during the day
2. It was a rainy day

Close-up of race-car engine
In the back of the garage is this engine from a racing car. Martin spent 4½ hours in order to replace the broken cylinders and repair the garage.

Close-up of infant on life-support
In the back seat of the car was this 10-month old boy. Martin operated for 4½ hours, resetting the infants’ broken hip and stabilising the collapsed lung.

Portrait of mechanic in front of garage
He is pleased that his son came to see him at the garage. It does not happen often that he works on such powerful engines.

Portrait of surgeon in a hallway
He is pleased that his son came to see him at the hospital. It does not happen often that he saves such young infants.

Mother walking away from a building
Victoria leaves the garage being late for her job. Phillip will stay the rest of the day with his dad.

Mother walking away from a building
Victoria leaves the hospital being late for her job. Phillip will stay the rest of the day with his dad.

Mother walking toward a phone booth
Heading to work, Victoria passes a police station. She realises that she is still carrying Phillip’s ball.

Mother walking toward a phone booth
Heading to work, Victoria passes a police station. She realises that she is still carrying Phillip’s ball.

Mother talking on the phone
She calls her boss to apologize for her delay. The boss tells Victoria to take the rest of the day off.

Mother talking on the phone
She calls her boss to apologize for her delay. The boss tells Victoria to take the rest of the day off.

Mother waiting at a bus stop
She tries to hail a cab home at the number 3 bus stop. She is glad that she does not have to go back to work.

Mother waiting at a bus stop
She tries to hail a cab home at the number 3 bus stop. She is glad that she does not have to go back to work.
3. Man puts flowers on grave  
4. Man kneels on ground  
5. Man says something  
6. Man kisses a picture  
7. The picture is on the grave  
8. A priest is walking around the cemetery  
9. Priest sees the man kneeling  
10. A man walks up behind the kneeling man

**Middle segment**

**Arousing version (13 details)**

1. Second man speaks to kneeler  
2. Kneeler turns his head  
3. Second man pulls gun from cloak  
4. Gun has silencer on it  
5. Second man shoots kneeler  
6. He shoots him one time  
7. Kneeler is wounded in chest or back  
8. Blood splatters on statue  
9. Kneeler falls onto grave  
10. Killer aims gun at fallen man’s head  
11. Priest says ‘For God’s sake, No!’  
12. Killer aims gun at priest  
13. Killer backs up

**Neutral version (11 details)**

1. Second man speaks to kneeler  
2. Kneeler turns his head  
3. Priest looks up to the sky  
4. Priest closes umbrella  
5. Priest looks on the ground  
6. Priest picks something up off ground  
7. Priest places object on top of grave  
8. Priest walks around a corner  
9. Puddles seen on pathways  
10. Priest shakes umbrella  
11. Close-up of statues on graves

**End segment – both versions (14 details)**
1. Second man/killer squats on ground
2. Second man/killers picks up something
3. Second man/killer backs slowly away
4. Second man/killer walks off
5. Priest goes to fallen man
6. Fallen man is not dead yet
7. Priest starts to pray
8. A (third) man is standing behind a nearby grave
9. The man is watching the priest
10. The man has blond hair
11. The man is wearing an overcoat
12. The man is wearing a tie
13. The man is wearing a white shirt
14. The man is Caucasian