# Journal of Autism and Developmental Disorders

**ERP Correlates of Recognition Memory in Autism Spectrum Disorder**

---Manuscript Draft---

<table>
<thead>
<tr>
<th>Manuscript Number:</th>
<th>JADD-D-12-00384R1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Title:</td>
<td>ERP Correlates of Recognition Memory in Autism Spectrum Disorder</td>
</tr>
<tr>
<td>Article Type:</td>
<td>Article</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Memory; autism spectrum disorder; event-related potential; Recognition; Old/New effect</td>
</tr>
<tr>
<td>Corresponding Author:</td>
<td>Esha Massand, PhD</td>
</tr>
<tr>
<td></td>
<td>UNITED STATES</td>
</tr>
<tr>
<td>Corresponding Author Secondary Information:</td>
<td></td>
</tr>
<tr>
<td>Corresponding Author's Institution:</td>
<td></td>
</tr>
<tr>
<td>Corresponding Author's Secondary Institution:</td>
<td></td>
</tr>
<tr>
<td>First Author:</td>
<td>Esha Massand, PhD</td>
</tr>
<tr>
<td>First Author Secondary Information:</td>
<td></td>
</tr>
<tr>
<td>Order of Authors:</td>
<td>Esha Massand, PhD</td>
</tr>
<tr>
<td>Dermot Mary Bowler, PhD</td>
<td></td>
</tr>
<tr>
<td>Laurent Mottron, PhD</td>
<td></td>
</tr>
<tr>
<td>Anthony Hosein</td>
<td></td>
</tr>
<tr>
<td>Boutheina Jemel, PhD</td>
<td></td>
</tr>
<tr>
<td>Order of Authors Secondary Information:</td>
<td></td>
</tr>
<tr>
<td>Abstract:</td>
<td>Recognition memory in Autism Spectrum Disorder (ASD) tends to be undiminished compared to that of typically developing (TD) individuals (Bowler, Gardiner and Gaigg, 2007), but it is still unknown whether memory in ASD relies on qualitatively similar or different neurophysiology. We sought to explore the neural activity underlying recognition by employing the old/new word repetition event-related potential (ERP) effect. Behavioural recognition performance was comparable across both groups, and demonstrated superior recognition for low frequency over high frequency words. However, the ASD group showed a parietal rather than anterior onset (300-500 ms), and diminished right frontal old/new effects (800-1500 ms) relative to TD individuals. This study shows that undiminished recognition performance results from a pattern of differing functional neurophysiology in ASD.</td>
</tr>
<tr>
<td>Response to Reviewers:</td>
<td>I have addressed each of the concerns below each point.</td>
</tr>
</tbody>
</table>

Reviewer #1: A single ERP experiment compared recognition memory with high and low frequency words compared adults with Autism Spectrum Disorder (ASD) to typically developing (TD) adults. Recognition memory accuracy did not differ between ASD and TD individuals, but ERP old/new effects showed some differences. The TD group should a typical pattern of 300-500 ms old/new effects over mid-frontal regions, followed by 500-800 ms old/new effects over parietal regions, followed by 800-1500 ms old/new effects over right frontal regions. The ASD group differed in having more a more parietal distribution of the early 300-500 ms effects, as well as having no discernable late 800-1500 ms effects.

Overall, I found the study to be well done, though I do have some questions about the methods below. The results seem reasonably solid and interesting. As such, assuming this is the first ERP study of recognition memory in ASD, I think a revision would merit publication. My main concern involves interpretation.
1. My main concern involves interpretation of the results.

A. There are some issues with the interpretation of the 300-500 ms mid-frontal old/new effects. The Introduction suggests, "an early midfrontal old/new effect (300-500 ms) indexing semantic memory judgements (Curran, 2000; Curran & Cleary, 2003; Curran & Dien, 2003)" (p. 4). It is odd that Curran's work is cited here because he (and most others, reviewed in Rugg & Curran, 2007). Think that these effects are related to the familiarity process that is described within dual-process theories of recognition memory, whereas the later parietal old/new effects reflect the second recollection process. The "semantic memory" perspective is probably closest to Voss and Paller's views. Somewhat relatedly, in reference to the discussion, I don't think anybody seriously agrees with Tulving's original ideas that "know" judgments reflect semantic memory. Again, they are most often interpreted as being related to familiarity-based recognition memory. These results need to be discussed more thoroughly in the context of prevailing views about recognition memory and ERP old/new effects. One interesting recent finding that may relate to the semantic memory ideas being advance is an experiment that directly compared 300-500 ms frontal old/new effects to 300-500 parietal semantic priming effects (Bridger, E. K., Bader, R., Kriukova, O., Unger, K., & Mecklinger, A. (2012). The FN400 is functionally distinct from the N400. NeuroImage.)

In the introduction both references are now cited (Curran’s work and that of Voss and Paller), and both familiarity and semantic memory are mentioned a both explanations now relevant for the discussion section of this paper.

We have acknowledged that the most prevailing view in the typically developing literature is in terms of the recollection/familiarity distinction. However, as the literature in ASD is entirely in terms of the Remember/Know, and, episodic/semantic memory distinction, and, because the manuscript is thought to be most relevant to ASD researchers, the episodic/semantic distinction was originally highlighted. The prevailing views in the typical literature are now explained within the text and considerable adjustments have been made throughout. This section has been added to the discussion (page 13):

These studies drew on the distinction made by Tulving (2002), between episodic and semantic memory systems, which to some extent mirrors a distinction made by Mandler (2008), who argues for the memory processes of familiarity and recollection (see Gardiner, 2008 for further discussion). According to Tulving, semantic memory allows the conscious experience of ‘knowing’ in the absence of recollecting specific contextual details about the study episode. Episodic memory on the other hand, is concerned with the storage of contextually rich, personally experienced and unique events and allows ‘remembering’. Although ERP-related studies of memory have used both the semantic/episodic and the familiarity/recollection distinctions, here, we discuss our findings in terms of the episodic/semantic distinction, simply because the majority of studies of memory in ASD have used this framework.

B. Similarly, the normal parietal old/new effects should be discussed in the context of the standard perspective that these effects are related to recollection.

The manuscript has been modified in line with the comment above.

This statement on page 18 puzzled me: "This finding of diminished later parietal effects in ASD is paradoxical, given the behavioural evidence of diminished Remember and preserved Know responses (Bowler et al., 2000a, b, 2007)." Did you mean to say "undiminished"? If so, I agree. It is paradoxical that parietal old/new effects are undiminished if remember responses are diminished and free recall if typically impaired in ASD because both "remembering" and free recall should tap into recollection processes. The discussion needs to be clearer on these points.

Yes this was indeed a typo (!) and has been amended - now page 14.

C. The Introduction explains the late 800-1500 ms ERP old/new effects as, "a sustained positive ERP potential over right frontal scalp sites, associated with post-
retrieval processes” (p. 4), but the discussion never discusses the lack of these effects in ASD from this post-retrieval process perspective.

A more detailed description of the late frontal effect and its relationship with source memory tasks/recollection has been added to the introduction. We have described the literature that links the right frontal to the experimental task. Page 2.

2. Questions about Methods

A. It is problematic that completely separate words lists were used as targets and lures. The same words should have been counterbalanced or randomly assigned to target/lure conditions across subjects. How do we know that the old/new effects do not just reflect item differences rather than memory? The words were matched on several characteristics, but they could differ in other unmeasured ways. This deserves some discussion. Perhaps that best argument the authors could use is that the TD old/new effects looked very similar to other published studies that used properly counterbalanced stimuli.

A footnote has been added to the manuscript on page 8 relating to reviewer 1’s comments about the frequency manipulation.

B. Why use fewer lures than targets? And how were the 25 lures/list divided into high and low frequency?

A paragraph has been added to the stimuli section on page 3.

C. What was the inter-trial timing?

During the test phase participants were instructed to fixate on a central point for 600 ms, which was replaced by the test word for 200 ms. The word was then replaced by a fixation cross for 1500 ms, followed by an old-new response prompt. The next trial began once the participant had made their response. This is now explained in the presentation section on page 4.

3. I assume that “Accuracy scores were corrected for guessing by subtracting the number of false alarms from the number of hits” (p. 9), should be “proportion” rather than “number” because more targets than lures were tested. Please clarify.

Proportion is the correct term.

4. In reporting the number of trials entering into the ERP analyses (p. 11), please also include the minimum number of trials/subject/condition. This should be no fewer than 15 or 20.

The minimum number of trials per subject across the Old/ New and High/ Low conditions was 18. This is elucidated on page 6.

5. The paper should somewhere mention how the present TD results compare with previous ERP studies of recognition memory that have manipulated word frequency. I believe that Rugg has done 2 or 3 word frequency studies. It would be good to show that the present TD results are consistent with past results.

The results section now addresses previous word frequency findings in the English language (Rugg and Doyle, 1992; Rugg, Cox, Doyle & Wells, 1995) as well as in German (Russier, Probst, Johannes & Munte, 2003), (page 7).

Reviewer #2: This manuscript contains a well-written account of a study conducted using event-related potentials (ERPs) to investigate the relationship between the processes supporting recognition memory in typically developing (TD) individuals and individuals with autism spectrum disorder (ASD). ERPs were acquired while individuals completed a recognition memory task for high and low frequency words. Discrimination was equivalent for the two groups. The ERP old/new effects were not equivalent. This outcome is consistent with the view that not entirely the same processes contributed to
the same degree to recognition memory for TD and ASD individuals. The authors quite rightly acknowledge that the data they have do not permit stronger functional claims than this, but this finding is in my view important and worthy of report.

I have several recommendations that I think would strengthen this manuscript. Recommendations 2 to 3 are the most important.

1. Differences due to frequency here are marginal at best. I recommend that this is simply acknowledged and that the primary analyses are restricted to data collapsed across this dimension. This would not, I think, take anything substantive away from the impact of the manuscript.

The section on High and Low Old-New effects has been removed, the manuscript acknowledges the findings (presented the data in a figure), and limits the primary analyses to the collapsed data (page 7).

2. I recommend that the ERPs elicited by misses are analysed and reported as they may provide useful information about function. It seems to me that there should be sufficient trial numbers to do this, and perhaps even to analyse ERPs elicited by false alarms (collapsed across frequency in both cases).

There were enough trials to analyse misses. Figure 4 has been added to show there were no visible group differences in the ERPs for misses. These data are not considered further (page 8).

To investigate further, a 2 word Frequency (High/Low) x 2 Group mixed Repeated Measures ANOVA revealed no significant difference in A’ (F (1,34)= 1.96, p = n.s.) sensitivity scores, or in B’ response bias (F (1,34)= 0.03, p = n.s.) between the two groups’ behavioral responses.

3. I struggled to determine when the analyses of difference waves were being reported, and could find very few outcomes that did not include or at least refer to analyses including the old/new factor. This factor no longer exists when analysing difference waves. I recommend that the authors include a separate section in which the outcomes of the analyses on the scaled difference scores are reported, or that they report the outcome of the rescaled data analyses immediately after the interaction term that might reflect changes in distribution (e.g. ‘This interaction term remained significant following rescaling (F(XX) etc .....)). It is critical that the reader can identify the key outcomes that licence claims about changes between distributions.

The difference waves were calculated for the visualisation of the old-new effect on the 2D topographical maps only. This is now detailed on page 5. Difference waves were not used for the ANOVAs because the authors wanted to retain the positivity/negativity of the old-new difference and not just the magnitude. The variables and levels (old, and, new) for the ANOVAs are detailed on page 7.

We have taken the reviewer’s advice and reported the outcome of the rescaled data (where interactions involve the variable ‘Region’) immediately after the interaction term that reflects changes in distribution. In the text the following (underlined) has been added:

In the mid-frontal old-new time-window, the ANOVA yielded a significant Old-New by Region by Group interaction(F(4,31) > 3.21, p < .03), this interaction remained significant after z-scaling the data (F(4,31) > 2.94, p < .04), which indicated that the old-new ERP effect in the 300-500 ms time-window had the typical frontal distribution in the TD group but had a parietal distribution in the ASD group (cf. Figure 2). In the 500-800 ms time-window, the old-new ERP effect was maximal over centro-parietal regions in both groups. Between 800 and 1500 ms, there was a long lasting positivity, which was larger in response to old than to new words over right frontal scalp regions (Old-New X Region X Laterality, F(4,31) > 6.22, p < .005). This interaction remained significant after z-scaling the data (F(4,31) > 5.18, p < .003).

4. I think that the nets are unhelpful. The scalp distributions are sufficient and as the main point point of the manuscript is about differences between distributions rather
than differences between amplitudes it might be worthwhile showing each map with individual maxima and minima rather than a common scale so that amplitude differences can be seen: the ERP plots themselves already show the amplitude differences to some degree.

The nets have been removed, and to reflect the minima and maxima and retain an easy visual group comparison, iso-contour lines have been added to the figures. This should be sufficient along with the ERP plots themselves to see the amplitude differences.
ERP Correlates of Recognition Memory in Autism Spectrum Disorder

Esha Massand¹
Dermot M Bowler¹
Laurent Mottron²
Anthony Hosein²
Boutheina Jemel²

¹City University London and ²Hôpital Rivière-des-Prairies, Université de Montréal, Montréal

Running Head: Recognition Old/New effect in ASD

Correspondence to: Esha Massand

Autism Research Group
Department of Psychology
City University
Northampton Square
London EC1V 0HB
UK
Tel: 0207 040 8380
Fax: 0207 040 8581
E-Mail: Esha.Massand.1@city.ac.uk
Abstract

Recognition memory in Autism Spectrum Disorder (ASD) tends to be undiminished compared to that of typically developing (TD) individuals (Bowler, Gardiner and Gaigg, 2007), but it is still unknown whether memory in ASD relies on qualitatively similar or different neurophysiology. We sought to explore the neural activity underlying recognition by employing the old/new word repetition event-related potential (ERP) effect. Behavioural recognition performance was comparable across both groups, and demonstrated superior recognition for low frequency over high frequency words. However, the ASD group showed a parietal rather than anterior onset (300-500 ms), and diminished right frontal old/new effects (800-1500 ms) relative to TD individuals. This study shows that undiminished recognition performance results from a pattern of differing functional neurophysiology in ASD.

Keywords: Memory, Autism Spectrum Disorder, Event-Related Potential, Recognition, Old/New effect

Corresponding author: Esha.Massand.1@city.ac.uk
One of the intriguing features of ASD is the profile of memory abilities often observed in this population. In general, HFA individuals show comparable immediate memory, cued recall (Boucher and Lewis, 1989) and recognition (Bowler, Gardiner & Gaigg, 2007) but impairments on measures of free recall. Some individuals with autism present exceptional memory performance in their domain of expertise (e.g. Mottron et al., 1998). They characteristically fail to use semantic relations amongst studied items to aid free recall (Bowler, Matthews and Gardiner, 1997; see Tager-Flusberg, 1991, for similar findings in LFA individuals). Although first noted by Boucher and Warrington (1976), this pattern of spared and impaired memory performance has been developed by Bowler and Gardiner into the Task-Support Hypothesis (TSH), (Bowler et al., 1997; 2004), which states that the performance of individuals with ASD will be closer to that of TD individuals if support is provided at test. Support is defined as any information relevant to the resolution of the task being present in the test materials. Supported test procedures include cued recall and recognition, where participants are presented with a clue (the studied word) and are asked whether or not it figured in the study list. In contrast, tasks such as free recall do not offer this kind of support, and often show diminished performance in individuals with ASD. This is especially the case when the task involves multiple trials (Bowler, Gaigg & Gardiner, 2008b; Bowler, Mottron & Limoges, 2009) or the use of semantic relations among studied items to enhance recall (Bowler, Matthews & Gardiner, 1997).

Confirming evidence for the TSH and for difficulties in processing inter-item relations has been provided by numerous subsequent studies. In one example, Bowler, Gaigg and Gardiner (2008a) tested the effects of item versus context relatedness on recall and recognition in adults with ASD by asking individuals to study a series of words presented on a screen, inside a red rectangle. Participants were asked to ignore context words (words that were either related or unrelated to the study words) that were presented outside the red rectangle. Participants then took part in a forced-choice recognition memory test that included both studied (inside the rectangle) and context (outside the rectangle) words. The results revealed that both groups recognised more target words than context words, and more related words than unrelated words. However despite this apparent typical memory performance, when the experiment was repeated with a recall rather than forced choice task, the recall of individuals with ASD, unlike that of the comparison participants, was not enhanced for related words. These findings show that contextual semantic relations are encoded sufficiently for people with ASD to benefit from them under recognition but not free recall test procedures.

Contextual details also play a role in the superior recognition by TD individuals of low frequency words (LFW) over high frequency words (HFW). According to Gardiner and Java, (1990) and Gardiner, Richardson-Klavohn and Ramponi (1997), the increased recognition rate for LFW in TD individuals, denotes an increase in the contextual detail with which these words are encoded during the study phase, which enhances recognition, recall and remembering. Word frequency manipulations have also been shown to enhance recognition in individuals with ASD (Mottron & Belleville, 1996; Bowler, Gardiner and Grice, 2000). Such observations lend further support for the view that similar underlying memory strategies may be used by TD and ASD individuals during recognition memory tasks (see Bowler et al., 2007). Yet, the
need for task support at retrieval suggests an underlying atypicality in the functioning of the memory system in individuals with ASD, which implies that when they perform memory tasks, they may be employing different underlying processes either at a cognitive or a neural level.

One way of studying neural processes underlying recognition memory uses the method of event-related potentials (ERPs, see Kutas & Federmeier, 2011; Rugg & Curran, 2007 for reviews). Many studies have demonstrated ERP differences between old (studied) and new (unstudied) items, such as words (Sanquist, Rohrbaugh, Syndulko & Lindsley, 1980; Rugg & Nagy, 1989), faces (Münte, Brack, Grootheer, Wieringa, Matzke, & Johannes, 1997; Guillem, Bicu, & Debruille, 2001) and objects (Van Petten, Senkfor, & Newberg, 2000). In general, ERPs are more positive in response to correctly recognised items relative to new items. This positive potential difference, referred to as Old-New ERP effect (Rugg, Cox, Doyle & Wells, 1995), occurs at ~300ms post stimulus and lasts several hundred milliseconds.

Further investigations of memory-related ERP indices identified three spatio-temporally distinct old-new ERP effects underlying specific memory retrieval processes (see Rugg & Curran, 2007 for a review). These ERP-memory related indices, although their respective memory-related functions are still controversial, encompass (1) an early mid-frontal old-new effect (300–500 ms) indexing familiarity judgements (Curran, 2000; Curran & Cleary, 2003; Curran & Dien, 2003), or semantic memory (Voss and Paller, 2008, 2009) and (2) a parietal, positive ERP effect (400–800 ms) related to recollection (i.e., the “parietal old-new effect”; Allan, Wilding, & Rugg, 1998; Curran, 2000; Curran & Cleary, 2003; Curran & Dien, 2003), and a sustained positive ERP potential over right frontal scalp sites, associated with retrieval processes, whenever a retrieval attempt is ambiguous, requires evaluation, such as in source memory tasks (Allan & Rugg, 1997, 1998; Allan, Dolan, Fletcher & Rugg, 2000; Curran, Schacter, Johnson & Spinks, 2001; Donaldson & Rugg, 1999; Johnson, Kounios & Norde 1996; Ranganath & Paller, 2000; Wilding & Rugg, 1997a,b; Wolk, Sen, Chong, Riis, McGinnis & Holcomb et al., 2009).

ERP methodology has proved useful in revealing atypical neural processes in the context of intact behavioural performance in various clinical groups. One such group that has particular relevance for ASD is the healthy ageing population. Memory in healthy aging, in common with ASD is characterised by better performance on supported tasks such as cued recall and recognition, (Craik & Anderson, 1999). Outside the memory domain, there are also similarities between healthy ageing and ASD, for example in terms of diminished ‘theory of mind’ (Slessor, Phillips & Bull, 2007) and impaired executive function (Buckner, 2004), making healthy ageing a useful heuristic for studying memory in ASD. Recognition memory in healthy ageing is reliably accompanied by significantly diminished Old-New word event-related potential (ERP) effects (Guillaume, Clochon, Denise, Rauchs, Guillery-Girard & Eustache et al., 2009).

The undiminished recognition memory performance and word frequency effect seen in individuals with ASD would lead us to predict typical ERP old-new effects in this group. However, the similarity between ASD and healthy ageing in the patterning of memory function across test procedures would lead us to predict different underlying neural processes in ASD, in particular, diminished old-new effects. In order to decide between these competing hypotheses we ran an ERP study using a standard recognition memory test with a word
frequency manipulation check. We tested a sample of ASD and age and IQ-matched TD individuals.

Method

Participants
Twenty-two participants with ASD (2 females) meeting DSM-IV-TR criteria (American Psychiatric Association, 2000) for ASD and 14 TD participants (2 females) took part in the study. ASD participants were of average FSIQ and matched closely to TD participants on gender, chronological age and FSIQ (as assessed by the Wechsler Adult Intelligence Scale III, Wechsler, 1997). Averages for age and IQ are presented in Table 1.

Participants were recruited from the Autism Specialized Clinic (Rivière-des-Prairies Hospital, Montreal). Diagnoses were based on the Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter & Le Couteur, 1994) and/or the Autism Diagnostic Observation Schedule (Lord, Risi, Lambrecht, Cook, Lenventhal & DiLavore et al., 2000).

TD participants were recruited from a panel maintained by the same institution, and all participants had normal or corrected to normal vision. Informed consent was obtained from each participant and the study was approved by the ethics committees at City University London and Hôpital Rivière-des-Prairies, Montreal.

Stimuli
The experimental stimuli were 300 target and 150 lure French words, individually extracted from the Desrochers & Bergeron (2000) corpus of 1,916 French nouns. Fewer targets than lures were used to reduce the time burden for participants, and, because the hit rate for lures was higher than for targets (thus proportions were equated). Three out of 6 blocks were comprised of 13 high frequency words, and 12 low frequency words, and the other three blocks comprised 13 low and 12 high frequency words.

Half target words and half lure words were of low frequency, (Targets: mean = 2.51 ± SEM = 0.12; Lures: mean = 2.37 ± SEM = 0.15; t(223) = 0.69, n.s.), and the other half were of high frequency (Targets: mean = 251.92 ± SEM = 13.31; Lures: Mean = 254.31 ± SEM = 23.1; t(223) = 0.096, n.s.) according to Baudot frequency index (Baudot, 1992). Low and high frequency target and lure words contained equivalent numbers of syllables (2.18 ± 0.07 and 2.13 ± 0.09 respectively for low frequency target and lure words, and 2.23 ± 0.07 and 2.36 ± 0.14 respectively for high frequency target and lure words; all t values < 0.68, n.s.) and letters (7.05 ± 0.16 and 7.01 ± 0.24 respectively for low frequency target and lure words, and 6.69 ± 1.56 and 6.85 ± 0.23 respectively for high frequency target and lure words; all t values < 0.91, n.s.). They were also equated with respect to imageability (the extent to which a word evokes a mental image) on the basis of Desrochers & Bergeron (2000) imagery norms (mean for all
words = 4.63 ± 0.07, all t values < 0.89, n.s.). Additional words with similar characteristics were used for the practice block.

**Presentation**

The experiment took place in a dimly lit and sound attenuated laboratory. Participants sat opposite a 17 inch computer screen, with their head restrained with a chin rest at a 60 cm viewing distance. Word presentation length varied from 2.5 cm to 11 cm and resulted in a minimum visual angle of 2.39° and a maximum viewing angle of 10.47°.

Participants completed one practice block and 6 experimental blocks, where each block comprised a study phase immediately followed by a test phase. The participants were instructed to study the words for a later memory test in which they were to decide whether the test word was presented in the study list (old-new response).

During the experimental blocks, each study list had 50 items (half LFW and half HFW) with an additional two buffer items (one buffer item at either end of the study list). At test, the 50 study items and twenty-five lures were presented. The blocks were presented in a random sequence and words were randomly ordered within blocks and balanced for orthographic/syllable length and imageability. During the study phase words appeared at a constant presentation rate of one every 2000 ms (1400 ms word presentation and 600 ms fixation on a central point measuring 0.8 cm). During the test phase participants were instructed to fixate on a central point for 600 ms, which was replaced by the test word for 200 ms. The word was then replaced by a fixation cross for 1500 ms, followed by an old-new response prompt. The next trial began once the participant had made their response.

**EEG/ERP Acquisition**

The electrical brain activity (electroencephalogram, EEG) was continuously recorded from 58 Ag/AgCl scalp electrodes mounted on an easy cap during the test phase. Electrode impedances were kept below 5kΩ. Bipolar electrooculogram (EOG) recordings were made using electrodes placed below and above the dominant eye (vertical EOG), and electrodes placed lateral to each eye (horizontal EOG). Signals from all electrodes were amplified with a bandpass from DC to 100 Hz, digitized at a 1024 Hz sampling rate and online referenced to the left earlobe. The right earlobe was actively recorded as an additional reference channel.

The data were processed using EEG Probe 4, a Linux ERP evaluation package (Nowagk & Pfeifer, 1996). To remain consistent with the majority of research to date (Düzel, Yonelinas, Mangun, Heinze & Tulving, 1997), the reference was changed offline to the average of the left and right earlobe recordings. Continuous EEG traces were band-pass filtered between 0.3-30 Hz. Prior to averaging EEG data associated with correct responses (hits and correct rejections) were examined for EOG and other artifacts using an automatic rejection procedure. EEG segments of 1900 ms durations (starting 200ms pre-word onset and lasting 1700 ms post-word onset) were rejected whenever the standard deviation in a 200 ms sliding time-interval exceeded 40 µV in EOG channels or 20 µV in any scalp electrode. Eye blinks were then corrected by subtracting from each electrode the PCA-transformed EOG components, weighted according to VEOG propagation factors (computed via linear regression).
To compute the ERPs, only artifact-free trials to old and new words associated with correct answers were used. Epochs of continuous EEG, including a 200 ms pre-stimulus and 1700 ms post-stimulus period, were averaged from each subject separately 1) for old and new words and 2) for old and new items separately for low and high frequency words. ERP difference waves were also computed to visualise the magnitude and topographical distribution of old-new effects in ASD and TD participants: 1) for all words (i.e., old minus new words) and 2) for high and low frequency words (i.e., old high minus new HFW, old low minus new LFW).

Data analysis

For the behavioural data, the proportion of hits (i.e., correct old responses to studied words) and false alarms (i.e., incorrect old responses to unstudied word) for each ASD and TD participant were calculated for all words and separately for low and high frequency words. Accuracy scores were corrected for guessing by subtracting the proportion of false alarms from the proportion of hits.

For the electrophysiological results, mean ERP amplitude measures were computed at each scalp electrode using three time-windows: 300-500 ms, 500-800 ms, and 800-1500 ms, encompassing the latency period of the mid-frontal old-new effect, the parietal old-new effect, and the late frontal old-new effect. These measurements were performed on the ERP averages for old and new words, both collapsed and separately for low and high frequency words.

A series of ANOVAs were conducted. To reduce Type 1 error as result of multiple comparisons, electrodes were clustered into five bilateral and three midline scalp regions of interest; left and right Fronto-Temporal (AF7/F7/F5/FT7/FC5, and AF8/F8/F6/FT8/FC6), left and right Frontal (AF3/F1/F3/FC1/FC3, and AF4/F2/F4/FC2/FC4), left and right Temporal (T7/C5/C3/TP7/CP5, and T8/C6/C4/TP8/CP6), left and right parietal (CP1/CP3/P3/P1, and CP4/CP2/P2/P4), left and right Occipito-Temporal (P7/P5/PO7/PO3, and P8/P6/PO8/PO4), and midline Frontal (Fpz/Fz/FCz), Central (C1/Cz/C2/CPz), and Parieto-Occipital regions (Pz/POz/O1/O2/Oz).

The magnitude and scalp distribution of Old-New ERP effects between groups were assessed with the ERP amplitudes. To ensure that topographic comparisons of ERPs were not confounded by differences in the magnitude of the Old-New effect, significant interactions involving Region by experimental, and/or by group factors were further investigated after vector-length normalization of the ERP amplitude measurements (McCarthy & Wood, 1985).

Results

Behavioural results

The raw data and corrected scores for the ASD and TD groups are presented in Table 2. The behavioural data were analysed using a 2 word Frequency (High/Low) x 2 Group mixed Repeated Measures ANOVA. There was no significant difference in the corrected recognition scores (F (1, 34) = 0.47, p = n.s., mean recognition in the TD group was 0.56 and ASD group
was 0.61). There was a main effect of word Frequency (F (1, 34) = 45.58, p < .01) where recognition was greater for LFW (M = 0.66) compared to HFW (M = 0.51). There was no word Frequency x Group interaction (F (1, 34) = 0.10, p = n.s.) demonstrating that both groups recognised more low frequency than high frequency words). Both groups recognised a similar proportion of false alarms (F (1, 34) = 0.004, p = n.s.), and both groups also made more false alarm judgements to HFW (M = 0.15, SD = 0.14) than LFW (M = 0.10, SD = 0.10), (F (1, 34) = 14.05, p < .01).

**ERP results**

The analyses of the electrophysiological data were conducted primarily to investigate potential differential old-new effects between ASD and TD groups. At this stage of analyses, trials corresponding to low and high frequency words were collapsed. The mean number of artifact-free trials with correct answers included for the ERP averaging of old and new words for ASD and TD groups, was 171.3 (S.D. = 49.7) and 156.8 (S.D. = 39.3) for old words, and 99.9 (S.D. = 29.6) and 102.8 (S.D. = 17.2) for new words respectively. The number of trials did not differ between the two groups (independent samples t-test (34) < .94; p > .3). Secondly, we examined old-new ERP effects with respect to word frequency. The ERP averages for these analyses also included equivalent number of trials for ASD and TD groups (independent samples t-test (34) < .95; p > 0.3). The mean number of averaged trials, for ASD and TD groups, was 78.7 (S.D. = 26.4) and 71.5 (S.D. = 20.8) for old HFW, 92.6 (S.D. = 25.1) and 84.9 (S.D. = 21.5) for old LFW, 48.2 (S.D. = 15.3) and 49.2 (S.D. = 9.4) for new HFW, and 51.7 (S.D. = 14.5) and 53.1 (S.D. = 9.8) for new LFW respectively. The minimum number of trials per subject across the Old/ New and High/ Low conditions was 18.

**Old/new ERP effects**

The ERP waveforms to old and new words are shown at nine electrode locations in Figure 1, separately for the ASD and TD groups. For both groups, there were consistent ERP amplitude differences between old and new words, with old words eliciting greater positive ERP voltages than new words, known as old-new ERP effect. Figure 2 displays a 2D scalp distribution of the old-new ERP differences in the 300-500 ms, 500-800 ms and 800-1500 ms for TD and ASD groups. These old-new effects appear remarkably similar to other published studies (Rugg & Doyle, 1992; Rugg et al., 1995; Rugg & Curran, 2007).

**Old-New ERP effects for low and high frequency words**
As can be seen in Figure 3, old-new ERP effects appear to be larger for low than for high frequency words (consistent with previous studies that have manipulated word frequency, Rugg and Doyle, 1992; Rugg, Cox, Doyle & Wells, 1995; Russler, Probst, Johannes & Munte, 2003). ANOVAs with Old-New, word Frequency, Region and Laterality as within-subjects factors and Group as a between-subjects factor yielded significant main effects of Old-New and of word Frequency in the 300-500 ms (mid-frontal old-new), 500-800 ms (parietal old-new), and 800-1500 ms (late frontal old-new) time-windows, for both lateral ($F(1,34) > 4.71, p < .04$) and midline regions ($F(1,34) > 5.45, p < .03$). These simple effects however did not interact with Group ($F(1,34) < 1.47, p > .2$). Moreover, the topographical distributions of these effects varied as a function of word frequency and group (cf. Figure 3). As there were no group differences, the primary analyses were restricted to data collapsed across high and low word frequency.

The old-new ERP effects in TD participants were characterized by three spatially and temporally different ERP effects (cf. Figure 2), consistent with the old-new ERP effects reported in previous studies (e.g., Curran, 2000; Duzel et al., 1997): a mid-frontally distributed old-new ERP effect in the 300-500 ms time-window, a parietal old-new ERP effect in the 500-800 ms time-window (parietal old-new), followed by a long-lasting right-frontal old-new ERP effect in the 800-1500 ms time window (late frontal old-new). The old-new ERP effects, albeit present in ASD participants displayed striking topographical differences (cf. Figure 2). More specifically, old-new effect in the ASD group had a left parietal distribution during the time-range of the mid-frontal old-new effect (300-500 ms). Furthermore, while the old-new effect between 500 and 800 ms was parietal in ASD participants, similar to TD participants, the late frontal old-new effect between 800-1500 ms was markedly reduced in the ASD group. These effects were statistically assessed using ANOVAs with Group (2 levels; TD, ASD) as a between-subjects factor and Old-New (2 levels; Old, New), Region (and laterality for measures computed on the 5 lateral regions) as within-subjects factors, performed separately in the 300-500 ms (mid-frontal old-new), 500-800 ms (parietal old-new), and 800-1500 ms (late frontal old-new) time-windows.

ANOVA results yielded significant Old-New effects in the three time-windows, for lateral ($F(1,34) > 10.68, p < .002$) and midline regions ($F(1,34) > 5.77, p < .025$).

In the mid-frontal old-new time-window, the ANOVA yielded a significant Old-New by Region by Group interaction ($F(4,31) > 3.21, p < .03$), this interaction remained significant after z-scaling the data ($F(4,31) > 2.94, p < .04$), which indicated that the old-new ERP effect in the 300-500 ms time-window had the typical frontal distribution in the TD group but had a parietal distribution in the ASD group (cf. Figure 2). In the 500-800 ms time-window, the old-new ERP effect was maximal over centro-parietal regions in both groups. Between 800 and 1500 ms, there was a long lasting positivity, which was larger in response to old than to new words over right frontal scalp regions (Old-New X Region X Laterality, $F(4,31) > 6.22, p < .005$). This interaction remained significant after z-scaling the data ($F(4,31) > 5.18, p < .003$). Although this interaction did not involve the factor Group, Figure 2 clearly shows that this late frontal effect was only present in the TD group. Post-hoc old-new comparisons with
Bonferroni corrections, computed for only the right frontal scalp region, shows a significant Old-New effect for TD \(t(13)= 2.91, p = .024\), and not for ASD \(t(21) = 1.8, p = .16\).

**Misses**

To enable comparisons for missed (studied but not recognised) words, the ERPs were compared to those for Old and New words (HFW and LFW collapsed) in each group. The data are presented in Figure 4. Number of miss ERP trials in the TD group: \(M = 71, SD = 28\); ASD group: \(M = 79, SD = 40\).

**INSERT FIGURE 4 HERE**

The ERPs for missed studied words were more negative than for New words in both groups, and there were no striking group differences in the ERPs.

**Discussion**

The aim of the present study was to test the hypothesis that despite comparable behavioural performance on recognition memory tests by individuals with ASD, the need for task support on memory tests might be reflected in atypical underlying neural activity at retrieval. Specifically, we investigated whether or not comparable recognition memory performance in ASD was associated with the same ERP old-new effects observed in typical individuals, that is, (1) an early mid-frontal old-new effect (300–500 ms), (2) a parietal, positive ERP effect (400–800 ms) and, (3) a sustained positive ERP potential over right frontal scalp sites. When participants were required to make old-new judgments of high and low frequency words in a recognition memory task, we replicated previous studies (Bowler et al., 2000; Bowler et al., 2007) by finding no behavioural difference between ASD and TD groups in the proportion of words correctly recognised. In addition, we found that both participant groups showed the well-established recognition memory advantage for low frequency words (Glanzer & Bowles, 1976; Guttentag & Carroll, 1994).

The undiminished behavioural performance of the ASD group was accompanied by group differences in ERP activity. Topographical differences in the Old-New effect were observed between groups from 300-500ms, where TD individuals, in addition to showing larger old-new effects for low frequency versus high frequency words (Rugg and Doyle, 1992; Rugg et al., 1995), showed an anterior onset (300-500ms) whereas the focus was at left posterior and parietal regions for individuals with ASD. In addition, the Old-New effect was attenuated at right frontal regions from 800-1500ms in the ASD group. We also observed a parietal focus of activity from 300-800ms in the ASD group, which contrasts with joint frontal and parietal activity for TD individuals\(^1\).

\(^1\) The old-new effects reported here are unlikely to reflect item differences due to the assignment of word stimuli into ‘Target’ and ‘Lure’ categories, given that the old-new effects in the TD group look very similar to other published studies (Rugg & Doyle, 1992; Rugg et al., 1995; Rugg & Curran, 2007).
These results show that even during supported recognition memory test conditions (according to the TSH, Bowler et al., 1997; 2004), individuals with ASD show neural activity that is different from that of TD individuals. Recognition memory of individuals with ASD stems from enhanced involvement of the semantic (indexed by ‘Know’ responses in a recognition memory test, see Tulving, 1985) and diminished involvement of the episodic memory systems (indexed by ‘Remember’ responses) (Bowler, Gardiner & Grice, 2000; Bowler, Gardiner & Gaigg, 2007). These studies drew on the distinction made by Tulving (2002), between episodic and semantic memory systems, which to some extent mirrors a distinction made by Mandler (2008), who argues for the memory processes of familiarity and recollection (see Gardiner, 2008 for further discussion). According to Tulving, semantic memory allows the conscious experience of ‘knowing’ in the absence of recollecting specific contextual details about the study episode. Episodic memory on the other hand, is concerned with the storage of contextually rich, personally experienced and unique events and allows ‘remembering’. Although ERP-related studies of memory have used both the semantic/episodic and the familiarity/recollection distinctions, here, we discuss our findings in terms of the episodic/semantic distinction, simply because the majority of studies of memory in ASD have used this framework. In TD individuals, the earlier on-setting mid-frontal Old-New effect is reliably found to accompany ‘Know’ judgements in Remember-Know tasks (Rugg & Curran 2007; Curran, 2000; Mecklinger, 2000) and the parietal old-new from 400-800ms (Voss & Paller 2008; Friedman & Johnson 2000; Paller & Kutas 1992) as well as late-onsetting right frontal positivity (Wilding & Rugg, 1996; Wolk, et al., 2009) have been reliably associated with episodic ‘Remember’ responses.

The topographical differences observed in the present study for ASD individuals in these time windows suggest that both the semantic and episodic memory systems engage partially different neural generators in ASD compared to those engaged by individuals with typical development. More specifically, the early mid-frontal Old-New effect observed for the TD group was absent for ASD individuals, who instead showed only a parietal focus, which provides evidence to suggest that semantic judgements may not be performed typically in ASD. Whereas parietal activity was observed for the ASD group from 400-800ms, activity for the TD group in this time period was more widespread, including anterior positivity. The fact that individuals with ASD show reduced ‘Remember’ responses in behavioural tests (Bowler et al, 2000, 2007) is also consistent with the attenuation of the old-new effect at right frontal regions reported here.

The pattern of between-group ERP differences just described suggests that rather than recognition being supported by early Old-New effects associated with knowing in TD, recognition in ASD seems to be exclusively associated with later parietal Old-New effects linked with remembering in TD. This finding of undiminished later parietal effects in ASD is paradoxical, given the behavioural evidence of diminished episodic and typical semantic responses (Bowler et al., 2000a, b, 2007). It implies that the neurophysiological correlates of semantic memory rather than episodic memory are compromised in ASD and is precisely the opposite of what has been concluded on the basis of behavioural evidence.

Conclusions
These findings show that comparable behavioural recognition in ASD is underpinned by qualitatively and quantitatively different patterns of neural activity, suggesting that recognition memory in this population may be driven by the operation of a single system rather than two systems as in typically developed individuals. Because the task used here was a supported one, the findings represent a conservative estimate of the memory-related neural differences in ASD. We can speculate that less supported test procedures, such as free recall or semantic cued recall, might reveal additional group differences in the ERPs for individuals with ASD. We plan to carry out further studies directly investigating the ERP correlates of episodic and semantic memory and direct manipulations of context effects in ASD and comparison participants.
Acknowledgements

The authors would like to thank the individuals who participated in the study. The first author was supported by a research studentship and international exchange scholarship from the Department of Psychology, City University London.
References


Figure 1: ERP traces to Old (in red) and New (in blue) words for TD ($N = 14$) and ASD ($N = 22$) groups shown at nine selected electrodes.
Figure 2: 2D scalp distributions of Old/New mean ERP amplitude differences (Old minus New words) in the 300-500 ms, 500-800 ms and 800-1500 ms for TD (N = 14) and ASD (N = 22) groups.
Figure 3: 2D scalp distributions of Old/New mean ERP amplitude differences (Old minus New words) for high and low frequency words in the 300-500 ms, 500-800 ms and 800-1500 ms for TD ($N = 14$) and ASD ($N = 22$) groups.
Figure 4: ERP traces to Old (red), New (blue) and Missed (green) words for TD ($N = 14$) and ASD ($N = 22$) groups shown at six selected electrodes.
Figure Captions

Figure 1. ERP traces to Old (in red) and New (in blue) words for TD ($N=14$) and ASD ($N=22$) groups shown at nine selected electrodes.

Figure 2. 2D scalp distributions of Old/New mean ERP amplitude differences (Old minus New words) in the 300-500 ms, 500-800 ms and 800-1500 ms for TD ($N=14$) and ASD ($N=22$) groups.

Figure 3. 2D scalp distributions of Old/New mean ERP amplitude differences (Old minus New words) for high and low frequency words in the 300-500 ms, 500-800 ms and 800-1500 ms for TD ($N=14$) and ASD ($N=22$) groups.

Figure 4. ERP traces to Old (red), New (blue) and Missed (green) words for TD ($N=14$) and ASD ($N=22$) groups shown at six selected electrodes.
<table>
<thead>
<tr>
<th></th>
<th>ASD (N=22)</th>
<th></th>
<th>TD (N=14)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Age (years)</td>
<td>25.72</td>
<td>4.76</td>
<td>23.85</td>
</tr>
<tr>
<td>VIQ</td>
<td>106.34</td>
<td>12.58</td>
<td>106.77</td>
</tr>
<tr>
<td>PIQ</td>
<td>102.45</td>
<td>13.10</td>
<td>96.69</td>
</tr>
<tr>
<td>FIQ</td>
<td>104.79</td>
<td>11.98</td>
<td>102.08</td>
</tr>
</tbody>
</table>

Table 1.

Mean and standard deviation for age and IQ measures (WAIS-III or WISC-III) for the TD and ASD groups. VIQ = Verbal IQ. PIQ = Performance IQ. FIQ = Full-scale IQ.)
Table 2

Mean and standard deviation of accuracy scores (proportions) for the TD and ASD groups are displayed separately for High and Low frequency words.

<table>
<thead>
<tr>
<th></th>
<th>TD (N=14)</th>
<th>ASD (N=22)</th>
<th>Both Groups (N=36)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>Hits</td>
<td>0.64</td>
<td>0.17</td>
<td>0.72</td>
</tr>
<tr>
<td>FA</td>
<td>0.15</td>
<td>0.13</td>
<td>0.10</td>
</tr>
<tr>
<td>Hits-FA</td>
<td>0.49</td>
<td>0.20</td>
<td>0.62</td>
</tr>
</tbody>
</table>
Author Note

Dr. Esha Massand, Department of Psychology, City University London, UK.
Professor Dermot M. Bowler, Department of Psychology, City University London, UK.
Professor Laurent Mottron, Université de Montréal, Montréal, Canada.
Anthony Hosein, Université de Montréal, Montréal, Canada.
Dr. Boutheina Jemel, Université de Montréal, Montréal, Canada.

Dr. Esha Massand is now at the Center on Human Development and Disability, University of Washington, Seattle, WA, USA.

Acknowledgements

The authors would like to thank the individuals who participated in the study. The first author was supported by a research studentship and international exchange scholarship from the Department of Psychology, City University, London.

Correspondence concerning this article should be addressed to Dr. E. Massand, esha.massand.1@city.ac.uk.