Do readers with autism make bridging inferences from world knowledge?

David Saldaña
University of Seville and Institute of Cognitive Neuroscience, University College London

Uta Frith
Institute of Cognitive Neuroscience, University College London

Address for manuscript correspondence:
David Saldaña
Psicología Evolutiva y de la Educación
Universidad de Sevilla
Avda. Camilo José Cela s/n
41018 Sevilla (Spain)
Tel.: 34954554334; Fax: 34954557642
E-mail: dsaldana@us.es
Abstract

Individuals with autism frequently show impairments in text reading comprehension. This is often attributed to poor ability to draw inferences during reading and to inadequate access to relevant knowledge. The present study tested this hypothesis by measuring the time taken to read the same question, relating either to physical or to social world knowledge, when it was either relevant or irrelevant to the bridging inference evoked by a preceding two-sentence vignette. Sixteen normally developing adolescents and 16 adolescents with autism were matched on word reading accuracy, chronological age and vocabulary, but differed significantly in text comprehension. A strong priming effect was found, robust over Ss and over items: participants read those questions faster that were relevant to the inference evoked by the vignette than those that were irrelevant, and no interaction with group membership or type of knowledge was found. This indicates that readers with autism, just like controls, were activating appropriate world knowledge primed by implicit inferences while reading the vignettes. Thus the comprehension problems in these readers cannot be attributed to an inability to make implicit inferences or to draw on relevant world knowledge. Instead we suggest that these problems must be sought at a higher level of text processing.

Keywords: Autism, text reading comprehension, poor comprehenders, bridging inferences, world knowledge
Inferences in poor comprehenders with ASD

Do readers with autism make bridging inferences from world knowledge?

In a simple sequence, such as *John had left his umbrella at home, He got very wet*, the reader's inference that *it rained* allows both sentences to make sense together. The knowledge that umbrellas protect people from the rain, and that you will get wet if you do not have a means of protection, are the basis for arriving at the implicit inference “*it rained*”. Bridging inferences require readers to make use of their previous world-knowledge even in such simple vignettes (Graesser, Singer, & Trabaso, 1994). One plausible way to account for comprehension difficulties are problems in drawing such inferences (Cain & Oakhill, 1999). A number of studies have investigated children who comprehend text poorly despite adequate word reading abilities (Cain, 1996; Nation, Adams, Bowyer-Craine, & Snowling, 1999; Oakhill, 1984; Yuill & Oakhill, 1991) and problems with inferencing ability have been repeatedly demonstrated in this research (Cain & Oakhill, 1999; Cain, Oakhill, Barnes, & Bryant, 2001; Oakhill, 1984).

A large proportion of readers with Autism Spectrum Disorders (ASD) present the discrepancy between text and word reading that defines poor comprehension (Nation, Clarke, Wright, & Williams, 2006; Nation & Norbury, 2005). Various experimental studies with individuals with ASD have highlighted limitations in text comprehension. Snowling and Frith (1986), Happé (1997), and Joliffe and Baron-Cohen (1999) found that children with ASD were less proficient at providing context-appropriate pronunciation of homographs than controls. Joliffe and Baron-Cohen (1999, 2000) and Norbury and Bishop (2002) found that individuals with ASD made more mistakes than control participants when answering multiple choice and open-ended format questions about a text they had just read. These questions had been
constructed in such a way that for the correct answer it was necessary to make inferences from the text and world knowledge.

Multiple choice and open-ended format tasks are “off-line” tests in the sense that they measure accuracy in the response to inferential questions presented after reading the text, i.e. well after the inferences have been made. We wanted to test the hypothesis that the comprehension difficulties in ASD are due to difficulties at an automatic level of text comprehension, which can only be tapped with more implicit tasks. In the present study, a paradigm originally devised by Singer (Singer & Halldorson, 1996; Singer, Halldorson, Lear, & Andrusiak, 1992) was used to tap on-line, or validate, the presence of bridging inferences, which are thought to be crucial for forming a mental model or representation of the text (Graesser et al., 1994). These inferences allow the reader to fill in information that is not explicitly mentioned in the text. The task involves the presentation of two-sentence vignettes followed by a general knowledge question, which is either primed or not primed by the implicit inference. Unlike the other techniques, the accuracy of the answer is of no interest, only the speed of reading and responding to the question. The question afterwards is not directly about the inference, but related to the knowledge that would be primed by the bridging inference. We thus measure an on-line effect of priming caused by the inference, which occurred while reading the vignette. In line with previous research, it was expected that readers with ASD would show difficulties in the automatic production of bridging inferences and thus show less of a priming effect.

At the same time, we wanted to explore the possible influence of knowledge on inferencing by contrasting vignettes that involved physical knowledge and others that involved social world knowledge. Inferencing requires both the information extracted from the text and the use of world knowledge. Lack of world knowledge can be a reason for poor performance, but
in previous research it has been found to be insufficient in explaining poor inferencing. Cain et al. (2001) presented skilled readers and poor comprehenders with texts that required them to learn knowledge that was novel for both groups, thus controlling for differences in the knowledge base. Differences in inferencing were still apparent between poor and skilled readers. However, this study concerned poor comprehenders without ASD. Readers with ASD would be expected to have more limited access to social world knowledge in line with their impaired social functioning: theory-of-mind and social interaction deficits are typically found in this population (Baron-Cohen, Tager-Flusberg, & Cohen, 2000) and could hinder the development of an adequate social knowledge-base, which might prevent successful bridging inferences when reading socially-related texts.

Method

Participants

Sixteen male readers with ASD and good word reading accuracy but poor text comprehension and sixteen normally developing readers, all adolescents, took part in the study. Both groups were matched for chronological age, standard word reading scores and receptive vocabulary on the British Picture Vocabulary Scale (Dunn, Dunn, & Whetton, 1997). All participants presented a minimum word reading age of 12 and chronological age of 11. Children with ASD showed a minimum difference of 10 standardized points between their word reading and text comprehension scores as measured by the WORD reading test (Rust, Golombok, & Trickey, 1993). This was the maximum discrepancy accepted for control participants’ scores. As a result, groups were significantly different in their text comprehension scores, \( t(30) = -4.45, p < .001 \) (see Table 1). The comprehension subtest of the WORD involved answering factual and inferential questions about the specific content of a short text.
ASD pupils were recruited from special schools or autism units within mainstream schools in Greater London, Surrey or Kent (United Kingdom), all of which are only accessed with educational statements of special needs that explicitly include an ASD-related diagnosis. Normally developing participants were recruited from mainstream secondary schools in Greater London, and no diagnosis of ASD or other disorders was reported. The only exception was one participant who had received learning support during the first years of primary school.

**Tasks and Materials**

Following the paradigm by Singer et al. (Singer & Halldorson, 1996; Singer et al., 1992) we constructed short vignettes from two sentences, which necessitated an inference. This was followed by a general knowledge question which was either primed or not primed by the inference. Unlike other techniques in the field of text comprehension, the accuracy of the answer was of no interest, only the speed of reading and responding to the question.

An example of an *experimental* triplet is:

1. *The Indians pushed the rocks off the cliff onto the cowboys.*

2. *The cowboys were badly injured.*

3. *Can rocks be large?*

Here the inference (the rocks hurt the cowboys) primes the question. This is labelled a *relevant* triplet.

The other half of the *experimental* triplets, termed *irrelevant* triplets, were carefully matched to the relevant triplets. The second sentence and the question were identical, but the first sentence was different. An example is:

1. *The Indians pushed the cowboys off the cliff onto the rocks.*
2. The cowboys were badly injured.

3. Can rocks be large?

Here the inference is that the cowboys were hurt due to the fall, and hence the question about the size of the rocks is not primed by the inference. Thus the question should be read and responded to slightly slower than when appearing in the first triplet. Such priming was found by Singer with normal adults readers (Singer & Halldorson, 1996; Singer et al., 1992). These authors hold that the presence of inferences is shown by the fact that subsequent questions are primed when relevant to the inference made, but not primed when irrelevant. Note that the preceding sentences themselves were carefully constructed so as to rule out the possibility that it was the sentences rather than the bridging inference that was responsible for any priming. To this end, the words used in the triplet pairs were as similar as possible.

Two types of phrases, defining two content conditions, were used. In one type, the knowledge required for the inferences was of physical content, whereas in the other it was of social content. A relevant triplet with social content was: Maria had never won a race before. The tears streamed down Maria's face. Can people cry because they are happy? The corresponding irrelevant triplet was: Maria had never lost a race before. The tears streamed down Maria's face. Can people cry because they are happy?

In addition to the experimental triplets, subjects were presented with filler triplets. These had the same format, half of them containing general knowledge questions, with a correct “no” response, and half specific questions relating to the first or second sentence, with correct “no” as well as “yes” responses. Examples are, The books were stacked on a cardboard box. The box broke. Can cardboard support heavy objects? and Only a few people took part in the race. They didn't all get to the end. Did a large crowd take part in the race? We only analysed reading time
Inferences in poor comprehenders with ASD

Data for the questions in the experimental triplets, where the answer was always “yes”. The inclusion of filler items insured that readers were attentive to the task and read all the sentences without a bias for yes or no responses.

For each condition, an initial pool of 20 relevant triplets was produced, and the first sentence of 10 randomly selected triplets was manipulated in order to make the question irrelevant to the inference required. This resulted in the first of the two lists. In the second list, the remaining triplets were converted from relevant to irrelevant and the other 10 triplets were left as relevant. Participants were assigned to one of these two lists, so that different children read each of the matched pairs of triplets. Each list contained 10 relevant triplets and 10 irrelevant triplets embedded in 32 filler triplets. The order of the items was the same in both parallel lists, determined by pseudorandom assignment, with no condition repeated on more than three consecutive trials.

Procedure

Participants were tested individually in a quiet room in their school. In a first session, they were tested for word reading, text reading comprehension and vocabulary. In two successive sessions, separated by at least one week, they were presented with the physical and social knowledge content phrases.

The experiments were run on a laptop computer, with the “p” and “q” keys as “yes” and “no” responses respectively. On each trial, the signal “Ready?” was displayed in the centre of the screen. When the subject pressed the spacebar, a fixation cross was presented for 500 ms. followed by the first phrase. When the participant finished reading the phrase, he/she pressed the spacebar again to see the second phrase, and again for the final question. The question was displayed until the subject responded, or for a maximum of 4 seconds. Responses that exceeded
this period were considered errors. Following a 1000 ms. inter-trial period, the next trial was initiated.

Subjects were led through two initial training trials where it was explained how to use the response keys. When they had understood the task, they commenced responding in the self-paced manner explained above. There was a 30 second rest period halfway through the list. At the break and on completing the task, the subjects were informed of the percentage of correct responses to increase their motivation.

**Results**

Reading times for all correctly answered questions in experimental triplets were submitted to mixed ANOVAs, treating alternatively subjects ($F_1$) and phrases ($F_2$) as the random variable, with list and group as between-subject variables and type of triplet (relevant vs. irrelevant) and phrase content (physical vs. social) as within-subject variables in the subject-random analysis.

The mean reading time for questions when placed in the context of relevant triplets was 317 milliseconds faster than the reading times when they were placed in the context of irrelevant items, which was significant both by subject, $F_1 (1,28) = 30.77, p < .001, MS_e = 104,462$, and by items, $F_2 (1,76) = 5.20, p = .025, MS_e = 661,904$. The interaction of triplet type by phrase content was non-significant, $F_1 (1, 28) = 2.23, p = .146, MS_e = 147,782, F_2 < 1$, indicating that the priming effect was equally strong in the social and physical content phrases.

The difference in the size of the priming between the ASD and control groups was very small (24 milliseconds), while each group showed significant priming effects of 305 and 329 milliseconds. The interaction of group by triplet type ($F_1$ and $F_2 < 1$) was accordingly non-significant. The interaction of group by content, $F_1 < 1, F_2 (1,76) = 2.76, p = .101, MS_e =
195,409, and the three-way interaction of content, group and triplet type ($F's < 1$) were also non-significant. Table 2 shows the response times for the questions in both groups for each content type.

**INSERT FIGURE 1 AND TABLE 2**

Overall accuracy for all subjects on questions in all experimental triplets was .94 ($SD = .07$). No difference in overall accuracy was found between the ASD and control groups ($M = .92, SD = .08$ and $M = .96, SD = .03$, respectively) ($U = 86.50, z = -1.598, p = .11$). There were no differences on the physical ($U = 105.5, z = -0.893, p = .37$) or social content questions when analysed separately either ($U = 85.5, z = -1.708, p = .08$) (see table 2 for accuracy in each content type). Although not analysed further, we assessed mean reading times for each sentence and found no differences between the ASD and control groups in the reading times for the first or second sentences or for the questions in either experiment ($F's < 1$). This was expected since the groups were matched on word reading ability.

**Discussion**

Our study replicated the so-called validation effect found by Singer (Singer & Halldorson, 1996; Singer et al., 1992) for bridging inferences with different materials and with adolescents rather than adults. Our participants, including those with autism, read and answered the same question more quickly when primed by an inference evoked by the preceding text, and more slowly when this was not the case. Differences were similar to those found by Singer and colleagues in their experiments, and were significant both in the by-subjects and by-item analyses. We found the same result regardless of whether the critical questions targeted physical or social world knowledge. We can thus conclude that our readers were activating the knowledge
necessary for bridging inferences and for this purpose were forming an appropriate representation of the text.

The finding that the validation effect was present in both the autistic and typically developing readers, without any apparent group interaction, is surprising given current views about the causes of poor text comprehension in readers with ASD. It is particularly striking since the present group of readers with ASD consisted of individuals who all had significant difficulties in text reading comprehension when tested on a standardised test. Our hypothesis that inferencing would be impaired at the automatic level was thus resoundingly disconfirmed. Thus the results lead us to ask new questions about previous studies exploring inferences in readers with ASD (Joliffe & Baron-Cohen, 1999; Norbury & Bishop, 2002). Participants of the ASD group in Joliffe and Baron-Cohen’s (1999) study, for example, were less accurate at selecting the correct sentence to go with a short vignette. In order to choose this sentence it would seem necessary to have made the correct bridging inference. Could it be that they did in fact make the correct inference at an automatic level, but then were confused about the choices they had to make? If so, we would predict that the correct sentence would be read faster when primed by the inference, even if it was not selected.

Previous experiments showing inferencing problems in readers with autism involved answering questions that were specifically related to the preceding text (Joliffe & Baron-Cohen, 1999; Norbury & Bishop, 2002), just like the test we had used in order to assess text reading comprehension in the present sample. This type of task, which requires explicit processing of information extracted from text, is poorly performed by children with ASD (Nation et al., 2006). Note that the present experiment did not rely on the accuracy of the answers. Instead the critical results came from the comparison across participants of time taken to answer the identical non-
inferential question depending on whether it occurred in the relevant or irrelevant condition. Whereas the question does require access to the knowledge-base, evidence of the inference is obtained from the faster response time: faster response times can only result from the prior implicit activation of the knowledge-base.

It is possible that the explicit measures of previous studies and the implicit ones used here are tapping different levels of textual representation (Kintsch, 1998). Bridging inferences are only one of the ways in which readers are able to build their representation of the text. Other aids, such as cohesion markers and pronoun identification, assist in elaborating the microstructure of the text, i.e., the representation of its basic propositions and their primary relationships. These processes are relatively automatic. The present results indicate that bridging inferences themselves are actually taking place in readers with ASD whose comprehension appears to be poor on standard off-line measures. It is therefore likely that these readers are able to construct implicit inferences, without necessarily using them to form higher-level representations. The argument that subjects with ASD find it difficult to relate the text base to the context of global text representation (Joliffe & Baron-Cohen, 1999) would be compatible with this possibility. Likewise, it is possible that differences in the demands to keep text in memory in the standard presentation of off-line inferencing tasks (e.g. Cain & Oakhill, 1999; Norbury & Bishop, 2002; Oakhill, 1982) would penalise the readers with ASD.

Can we rule out that poorer accessibility of the relevant information in the knowledge base contributed to problems in text comprehension in the standard comprehension test? Our experimental stimuli were designed to explore this question, by allowing the comparison of physical and social knowledge. Kaland et al. (2005), following up on Happé’s previous work (1994), recently found evidence of poorer ability to answer inferential questions after reading a
short text when it referred to the mental states of the characters than when it was related to physical events, in an off-line comprehension task. The present results which show a strong validation effect in the readers with ASD suggest that the retrieval of social information was just as easy as that of physical information and just as efficient as that for typically developing readers. These results therefore suggest that the impairment found previously may not be due to poorer access to the knowledge base, but to poorer ability to integrate specific knowledge explicitly with the global text.

The possibility of intact automatic processing at the sentence level in individuals with impaired comprehension is consistent with recent results obtained in another on-line test using the measure of anticipatory eye movements while listening to sentences (Nation, Marshall, & Altman, 2003). Participants were asked to select an object from a four pictures when they heard it mentioned in a phrase. Experimental phrases included verbs that allowed anticipation of the target object before it was mentioned. Both poor comprehenders and control participants made equally fast anticipatory eye movements toward the targets. While the poor comprehenders presented different fixation patterns overall, their anticipative behaviour indicates that they were using the information activated while listening to sentences to guide on-line processing in the same way as normal comprehenders.

Some studies have shown that non-autistic poor comprehenders are less efficient at suppression, rather than activation, of information (Cain, 2006; De Beni & Palladino, 2000; De Beni, Palladino, Pazzaglia, & Cornoldi, 1998; Gernsbacher & Faust, 1991). Nation et al. (2003) hypothesized that this could be a possible explanation for the greater number and shorter duration of fixations to target objects of their poor comprehenders. Norbury and Bishop (2002) extended this argument to readers with ASD by suggesting that inferencing errors could reflect
difficulties in suppressing irrelevant information, which might limit access to new information that would be relevant to the text being read. Incorrect, but maybe more familiar, representations might instead be activated by the text, leading to comprehension errors and incorrect inferences. It is possible that this could occur at a situation model level or in building a mental model of the text as a whole. From our results it would seem less likely to be happening at the level of bridging inferences. If this were the case, knowledge brought to bear in the validation effect would include both the relevant and irrelevant information, or randomly one or the other, thus cancelling out a priming advantage.

Finally, the remarkable lack of differences in our study between the readers with autism and controls, cannot hide the fact that these participants have problems in comprehending text. However, our study has ruled out two more obvious sources for the problem: difficulty at an automatic level of inferencing and difficulty in accessing knowledge. This clears the way for a renewed effort to explain why children with autism who read very well, and can make bridging inferences, fail to comprehend what they are reading.
References


Inferences in poor comprehenders with ASD


Author Note

This study was possible thanks to grant number EX2004-0098 of the Postdoctoral Programme of the Spanish Ministry of Science and Education to the first author. We would like to thank the participating pupils and their parents and teachers, Sarah White, for her always useful help and input in different phases of this study, and Ade Owolabi, Sarah Johnson, Megan Fisher and Gemma Northam for their assistance in collecting the data.
Footnotes

1. Both conditions were initially designed as different experiments and were therefore not counterbalanced across subjects. However, since both experiments had equal design and participants and equivalent stimuli, they were analyzed jointly and are presented as different conditions for greater clarity. The complete original sets of phrases are available from the first author at dsaldana@us.es.

2. Data were not trimmed or truncated in any way, and analyses were performed on all correct items following recommendations by Ulrich and Miller (1994). However, analyses were also carried out with exclusion of outlier responses above and below 2 SD of individual subject by condition means and with medians. Results in both cases were similar to those reported here.
### Table 1

**Participant characteristics (mean with SD in parenthesis, range below)**

<table>
<thead>
<tr>
<th>Group</th>
<th>Age</th>
<th>Word Reading Standardized Scores</th>
<th>Text Reading Standardized Scores</th>
<th>BPVS Standardized Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASD</td>
<td>14:9 (1:11)</td>
<td>101.8 (9.6)</td>
<td>80.4 (13.9)</td>
<td>98.2 (23.9)</td>
</tr>
<tr>
<td>(n = 16)</td>
<td>12:4 – 19:4</td>
<td>86 – 120</td>
<td>57 – 104</td>
<td>53 – 147</td>
</tr>
<tr>
<td>Control</td>
<td>13:11 (1:11)</td>
<td>101.6 (8.0)</td>
<td>98.9 (9.1)</td>
<td>98.7 (14.0)</td>
</tr>
<tr>
<td>(n = 16)</td>
<td>11:6 – 18:8</td>
<td>81 – 116</td>
<td>78 – 114</td>
<td>79 – 134</td>
</tr>
</tbody>
</table>
Table 2

Mean reading times in milliseconds ± standard deviations (accuracy) by condition for correctly responded questions in experiments 1 (physical) and 2 (social)

<table>
<thead>
<tr>
<th>Group</th>
<th>Physical Knowledge Inferences</th>
<th>Social Knowledge Inferences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Irrelevant</td>
<td>Relevant</td>
</tr>
<tr>
<td></td>
<td>items</td>
<td>Items</td>
</tr>
<tr>
<td>ASD</td>
<td>2838 ± 718</td>
<td>2552 ± 744</td>
</tr>
<tr>
<td></td>
<td>(0.91)</td>
<td>(0.94)</td>
</tr>
<tr>
<td>Control</td>
<td>2523 ± 605</td>
<td>2377 ± 663</td>
</tr>
<tr>
<td></td>
<td>(0.96)</td>
<td>(0.95)</td>
</tr>
</tbody>
</table>