Journal of Neurolinguistics 45 (2018) 36–44

Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

Perceptual simulations during sentence comprehension: A comparison between typical adolescents and adolescents with autism spectrum disorder

Orna Peleg [a, b, *], Rotem Ozer [c], Tal Norman [a], Osnat Segal [c]

[a] The Program of Cognitive Studies of Language and its Uses, Israel
[b] Sagol School of Neuroscience, Tel-Aviv University, Israel
[c] Department of Communication Disorders, Sackler Faculty of Medicine, Tel-Aviv University, Israel

A R T I C L E I N F O

Article history:
Received 26 December 2016
Received in revised form 22 August 2017
Accepted 30 August 2017

A B S T R A C T

Consistent with embodied theories of language comprehension, several studies have shown that comprehenders automatically activate perceptual (visual) information of verbally described objects, even when this information in neither explicitly mentioned nor necessary to perform the task. To clarify the role of perceptual activations in meaning construal; and to identify a potential cause of comprehension difficulties in autism spectrum disorder (ASD), the present study examined the extent to which individuals with ASD activate perceptual representations during sentence comprehension. 16 adolescents with ASD and 16 typically developing (TD) controls were asked to decide whether an object depicted in a line drawing had been mentioned in a preceding sentence. In the *match condition*, the shape or the orientation of the object matched the one implied by the preceding sentence. In the *mismatched condition*, the shape or the orientation of the object did not match the one implied by the sentence. TD adolescents responded faster in the match than in the mismatch condition. In contrast, adolescents with ASD did not distinguish between the two conditions. Thus, compared to controls, ASD adolescents were less able to spontaneously activate perceptual information as a function of sentence context. The implications of these results are discussed.

1. Introduction

Embodied theories of language processing hold that sentences are understood by mentally simulating the state-of-affairs described by the sentence. That is, the same mental representations that are activated when we perceive real events are also activated in response to verbally described events. Language comprehension therefore involves not only the activation of linguistic representations and structures, but also the activation of sensorimotor representations associated with the described objects and events (e.g., Barsalou, 2008; Zwaan et al., 2004). In line with this proposal, several studies have shown that comprehenders automatically activate perceptual (visual) information of verbally described objects even when this information is not explicitly stated, but merely implied by the described situation (e.g., Stanfield & Zwaan, 2001; Zwaan, Stanfield, & Yaxley, 2002; Engelen, Bouwmeester, de Bruin, & Zwaan, 2011).

* Corresponding author. The Program of Cognitive Studies of Language and its Uses, Tel Aviv University, Tel Aviv 69978, Israel.
E-mail address: pelegor@post.tau.ac.il (O. Peleg).

http://dx.doi.org/10.1016/j.jneuroling.2017.08.003
0911-6044/

In these studies, participants are asked to read (or listen to) sentences describing objects in particular locations. Importantly, the object's visual shape (or orientation) changes as a function of its location. For example, the sentence "Ron saw the balloon in the air" implies an inflated balloon, whereas the sentence "Ron saw the balloon in the pack" implies a deflated balloon. Similarly, the sentence "Dan saw the toothbrush in the cup" implies a vertical toothbrush, while "Dan saw the toothbrush in the sink" implies a horizontal toothbrush.

To investigate whether comprehenders spontaneously activate such subtle perceptual details, a sentence-picture verification task is typically used, in which participants are asked to decide whether a depicted object (e.g., balloon) was mentioned in the preceding sentence (e.g., "Ron saw the balloon in the air"). The critical trials are always positive (i.e., the object was indeed mentioned), however, the object's shape (or orientation) may be consistent (e.g., an inflated balloon) or inconsistent (e.g., a deflated balloon) with the one implied by the sentence. In all these studies, responses were faster in the match than in the mismatch condition, suggesting that the mental representation resulting from sentence comprehension incorporates perceptual (visual) information associated with the described scene, even when this information in neither explicitly mentioned, nor necessary to perform the task.

These visual effects (i.e., faster responses in the match than in the mismatch condition) were obtained regardless of the specific task that was used (e.g., Madden & Zwaan, 2006), or the participants' age (e.g., Dijkstra, Yaxley, Madden, & Zwaan, 2004; Engelen et al., 2011). For example, using the same sentence-picture verification task, Engelen et al. (2011) demonstrated that even children as young as seven are able to activate and integrate perceptual information, as a function of sentence context. In that study, similar perceptual effects were found among children aged 7–13 years, both after listening to sentences and after reading them aloud. Importantly, the effect size did not increase as a function of age, suggesting that both younger and older children construct a perceptual simulation of the described object's shape and orientation during sentence comprehension. Such early visual effects are consistent with the idea that perceptual simulation of verbally described situations is obligatory and relatively effortless.

Nevertheless, although the interaction between linguistic and perceptual representations, is now well established, exactly what role these representations play in language comprehension is still under investigation (e.g., Mahon & Caramazza, 2008). In an attempt to establish a causal relationship between perceptual simulations and language comprehension, Madden and Zwaan (2006), examined the ability of low- and high-span comprehenders to activate implied perceptual representations of verbally describe objects. High-span comprehenders activated implied perceptual information faster than low-span comprehenders, suggesting that the strength of the links between linguistic and perceptual representations may influence the process of meaning construal. In particular, Madden and Zwaan (2006) have suggested that perceptual representations may act as a mechanism for construing the contextually appropriate sense of a word during language processing.

To further test the relationship between perceptual simulations and language comprehension, the present study focused on individuals with autism spectrum disorder (ASD). As detailed below, constructing contextually appropriate meanings can be a source of particular difficulty for individuals with ASD. If perceptual simulations are necessary for meaning construal, then comprehension difficulties in individuals with ASD may be associated with weaker connections between linguistic and perceptual representations. Thus, to clarify the role of perceptual simulations in language processing; and to identify a potential cause of comprehension difficulties in ASD, the goal of the current study was to investigate the extent to which individuals with ASD spontaneously activate perceptual information as a function of sentence context.

ASD is a neurodevelopmental condition characterized by impaired social interaction and communication; and restricted, repetitive patterns of behavior (DSM-V). Although both language and perceptual (visual) deficits are not part of the diagnostic criteria, individuals with ASD often exhibit impairments (or atypical behavior) in both domains. In particular, it has been claimed that individuals with ASD have a cognitive bias towards local rather than global processing (e.g., Frith, 1989; Happé & Frith, 2006). That is, they tend to focus on the 'trees' (i.e., the details, parts) rather than the 'forest' (i.e., the whole). In the visual domain, this cognitive bias is demonstrated by superior performance in tasks requiring local processing (e.g., O'Riordan, Plaisted, Driver, & Baron-Cohen, 2001; Plaisted, O'Riordan, & Baron-Cohen, 1998), together with impairments in tasks requiring global processing (e.g., Brosnan, Scott, Fox & Pye, 2004; Happe, 1996).

In the linguistic domain, support for reduced global processing comes from studies demonstrating impaired ability to use context in order to generate inferences (e.g., Norbury & Bishop, 2002) or in order to resolve ambiguous words (e.g.,Frith & Snowling, 1983; Happe, 1997; Jolliffe & Baron-Cohen, 1999). For example, Frith and Snowling (1983) asked children with and without ASD to read aloud sentences that contained homographs (words with the same spelling, but different pronunciation and meaning, such as *tear*). The sentences were biased towards the homograph's dominant ("In her eyes there was a *tear*") or subordinate ("In her dress there was a *tear*") meaning. In contrast to typically developing (TD) children who tended to read the contextually appropriate word, children with ASD, tended to choose the dominant pronunciation, irrespective of the sentential context. However, supporting the idea of a local bias rather than a global processing impairment (e.g., Happé & Frith, 2006), there is also evidence that individuals with ASD have the ability to process information globally if they are explicitly instructed to do so (e.g., Hadad & Ziv, 2015; Snowling & Frith, 1986). Thus, evidence from studies in both the visual and the linguistic domains suggest that in situations where TD individuals spontaneously process incoming information globally and in context, individuals with ASD tend to process information "piece by piece" (but see, Giora, Gazal, Goldstein, Fein, & Stringaris, 2012; Hahn, Snedeker, & Rabagliati, 2015, for different results with high functioning autistic individuals).

At the neurological level, this cognitive bias towards local rather than global processing in individuals with ASD, was recently explained in terms of "disrupted cortical connectivity" (e.g., Brock, Brown, Boucher, & Rippon, 2002; Just, Cherkassky, Keller, & Minshew, 2004; Just, Keller, Malave, Kana, & Varma, 2012; Kana, Keller, Cherkassky, Minshew, & Just, 2006). In particular, their superior performance on a variety of perceptual-visual tasks requiring local processing (e.g., O'Riordan et al., 2001; Plaisted et al., 1998) was found to be correlated with abnormal over-connectivity in posterior areas associated with visual processing (e.g. Kana et al., 2006; Mizuno, Villalobos, Davies, Dahl, & Müller, 2006). At same time, their integration difficulties were correlated with both abnormal connectivity in frontal areas responsible for higher level conceptual processes, and reduced connectivity across neural systems involved in complex cognitive tasks, such as sentence comprehension (e.g., reduced connectivity between posterior and frontal brain areas). For example, in a fMRI study conducted by Just et al. (2004), brain activations were measured while people with and without ASD comprehended sentences. Results indicated that degree of synchronization between the various participating cortical areas was consistently lower for the autistic than the control participants.

Nevertheless, although there is evidence that individuals with ASD exhibit cognitive differences at both the perceptual and the linguistic level; and although there is evidence for reduced connectivity between lower-level perceptual mechanisms and higher-level conceptual mechanisms in ASD, the relationship between linguistic representations (e.g., words or sentences) and perceptual representations (e.g., objects or events) among individuals with ASD, is relatively unexplored. The present study therefore examined the extent to which sentence comprehension in individuals with ASD involves the activation of perceptual representations associated with the described objects and events. In particular, using the sentence-picture verification task described above, we examined the ability of adolescents with and without ASD to activate and integrate perceptual information, as a function of sentence context. Given evidence that individuals with ASD exhibit a cognitive local bias at both the perceptual and the linguistic domains; and given evidence for reduced functional connectivity across different centers of the brain in this population, we expected perceptual effects (i.e., faster responses in the match than in the mismatch condition) in the TD group, but not in the ASD group.

2. Methods

2.1. Participants

Sixteen adolescents with ASD (11 males and 5 females; mean age = 17.8, SD = 1.3) and sixteen TD controls matched for age and gender (11 males and 5 female; mean age = 17.5, SD = 1.8) participated in the study. All participants were native speakers of Hebrew with no hearing or vision impairments. The ASD participants were assessed by psychiatric specialists and were diagnosed according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). They were all from the same special school for adolescents with ASD. The TD controls attended regular high schools in the same area. To assess non-verbal and verbal abilities, all participants were asked to complete, in addition to the experimental task, a non-verbal IQ test (Raven, Raven, & Court, 1998) and a verbal expressive vocabulary test (Kavé, 2005). In both tests, ASD adolescents' scores were significantly lower than their TD controls (Raven non-verbal IQ - ASD: M = 93.75; SD = 33.34, range: 55–145; TD: M = 137.18, SD = 11.5, range: 110–145; $p < 0.001$; Kave's expressive vocabulary test - ASD: M = 79.6; SD = 13.18, range: 52.1–95.8; TD: M = 95.96, SD = 2.18, range:89.6–97.9; $p < 0.001$). A more detailed inspection of the ASD adolescents' scores revealed not only large individual differences in verbal and non-verbal abilities (where some scored low while others scored high), but also large discrepancies between these two abilities, (where some scored high on the non-verbal test and low on the verbal test, and vice versa). This discrepancy between verbal ability and non-verbal ability made it impossible for us to match the two samples (ASD vs. TD) on both abilities. Thus, the two samples were matched for age and gender but not for general cognitive ability. As detailed below, these differences in cognitive ability were taken into account in the analysis of the results. In addition, because non-perceptual contextual difficulties (e.g., deficits in sentence-level semantic integration) may modulate the perceptual effect examined in this study, the ASD group also performed a semantic judgment task in which we tested their ability to activate and integrate semantic information as a function of sentence context.

2.2. Main experiment - the sentence-picture verification task

Materials: The experimental stimuli consisted of 40 sentences taken, with a few changes, from the Engelen et al. (2011) study. From their larger list of Dutch sentences, we selected those that described objects and situations that Israelis are familiar with, and that we could easily translate into Hebrew. All the sentences had the same structure: “X saw a particular object in a particular location”. The sentences were organized in pairs such that each pair described the same object (e.g., balloon) in two different locations (e.g., “in the pack” or “in the air”). Importantly, these two different locations implied two different shapes/orientations of the described object. For example, “Ron saw the balloon in the pack” implied a deflated balloon, while “Ron saw the balloon in the air” implied an inflated balloon. Similarly, “Dan saw the toothbrush in the cup” implied a vertical toothbrush, while “Dan saw the toothbrush in the sink” implied a horizontal toothbrush.

The 40 experimental sentences were accompanied by 40 black and white line drawings, taken from Stanfield and Zwaan (2001); Zwaan et al. (2002); Zwaan and Pecher (2012); https://osf.io/t6mrd/). Thus, each experimental pair (e.g., “Ron saw the balloon in the pack/air”) was associated with two pictures (e.g., one showing a deflated balloon and the other one showing an inflated balloon). To ensure that our translated items can evoke the appropriate visual representation, and that Hebrew speakers can recognize the pictorial objects, a pretest was conducted. In the pretest, 20 native speakers of Hebrew that did not participate in the experiment itself, were presented with a list of sentences and their corresponding pictures. After each sentence, two pictures of the described object were presented – one picture matched the implied shape or orientation, while the other one did not. Participants were instructed to choose the picture that best matches the situation that the sentence describes. Two lists were created, such that each list contained only one member of each sentence pair presented with the same two pictures. Only items that received 100% agreement were included in our final set, which as mentioned above, consisted of 40 sentences and 40 pictures.

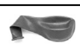
In order to create the two experimental conditions, each sentence was presented in either a *match* or a *mismatch* condition (see Table 1). In the *match condition*, the shape or the orientation of the object matched the shape or the orientation implied by the sentence (e.g., the sentence “Ron saw the balloon in the air” was paired with a picture of an inflated balloon). In the *mismatch condition*, the shape or the orientation of the object did not match the shape or the orientation implied by the sentence. (e.g., the sentence “Ron saw the balloon in the air” was paired with a picture of a deflated balloon).

The 80 picture-sentence combinations (40 sentences X 2 pictures) were divided into 4 lists (2 sentences X 2 pictures, for each object). Each list included only one sentence-picture combination of each object. In addition to the 20 sentence-picture experimental items, each list contained 40 sentence-picture items that served as fillers. Ten items presented a picture of the described location, rather than the described object. This was done in order to prevent participants from merely paying attention to the object. The other 30 items presented pictures of objects that were not mentioned in the preceding sentence, and thus required a “no” response. This was done in order to equate the number of ''yes'' and ''no'' responses, as both the 20 experimental items and the 10 location-filler items presented pictures of objects that were mentioned in the preceding sentence, and thus required a “yes” response.

Each participant completed two lists. Thus, each participant saw two sentence-picture combinations of each experimental item, but each time with a different sentence version (Ron saw the balloon in the pack/air) and a different picture version (an inflated or a deflated balloon). Similarly, the fillers in each list described the same objects in different locations (Golan saw the baby in the carriage/swing) and were paired with different pictures. Overall, each participant saw 40 experimental items (20 in the match condition and 20 in the mismatch condition) - requiring a “yes” response; 20 location-fillers also requiring a “yes” response; and 60 unrelated fillers, requiring a “no” response.

Apparatus: The experiment was constructed and run using E-prime software version 10.242. Sentences were presented on the screen in black letters on a white background, in Times New Romans Hebrew font, size 30. Pictures, fitted to occupy a 3 × 3

Table 1
Translated examples of stimuli.

Condition	Sentence	Picture	Correct Response
Match (shape)	Ron saw the balloon in the pack		Yes
Match (shape)	Ron saw the balloon in the air		Yes
Mismatch (Shape)	Ron saw the balloon in the pack		Yes
Mismatch (Shape)	Ron saw the balloon in the air		Yes
Match (Orientation)	Dan saw the toothbrush in the cup		Yes
Match (Orientation)	Dan saw the toothbrush in the sink		Yes
Mismatch (Orientation)	Dan saw the toothbrush in the cup		Yes
Mismatch (Orientation)	Dan saw the toothbrush in the sink		Yes

inch square surrounded by 1.5-inch white frame (total size of framed pictures is 6 × 6 inch), which was presented centrally on a gray background.

Procedure: Participants were tested individually in a quiet room. They were asked to read sentences and then to verify whether or not a depicted object was mentioned in the preceding sentence. At the beginning of the experiment, participants were given instructions and two examples, followed by 10 practice trials consisting of 5 related and 5 unrelated items. All trials had the same sequence of events (see Fig. 1). At the start of each trial, participants were presented with a central fixation marker for 500 ms. The offset of the marker was followed by a sentence, which was also read aloud by the experimenter. The sentence remained on the screen for 3000 ms, and was then replaced by a central fixation point. After 1000 ms, this point was replaced by the picture, which remained on the screen until the participant responded.[1] Response latencies were measured from the onset of the picture presentation. The trials were presented in random order. Each list of 60 trials - 20 experimental items and 40 fillers - took approximately 15 min to complete. Each participant completed two lists, separated by a 5 min break.

2.3. Additional experiment for ASD participants – the semantic judgment task

In order to control for non-perceptual contextual difficulties that might modulate the perceptual effect examined in this study, the ASD group also performed a semantic judgment task, in which we tested their ability to activate and integrate semantic information as a function of sentence context. Specifically, they were asked to judge whether or not a target is semantically related to its preceding prime.

The experimental materials were all taken from Peleg and Eviatar's (2008) study. The critical stimuli were 23 Hebrew prime-target pairs comprising an ambiguous prime (e.g., bank) and a target related to its subordinate meaning (e.g., river). Each prime-target pair was presented in two conditions: a "no-context" condition and a "context" condition. In the "no-context" condition, the prime was presented without context (e.g., bank - river). In the "context" condition, the same prime was presented after a subordinate biasing context (e.g., The fisherman sat on the - bank - river). Participants performed the task once without context and once with context. In both cases the expected answer is "yes". However, if participants are able to use context in order to enhance the activation of the contextually relevant meaning, then responses should be more accurate in the "context" condition than in the "no context" condition.

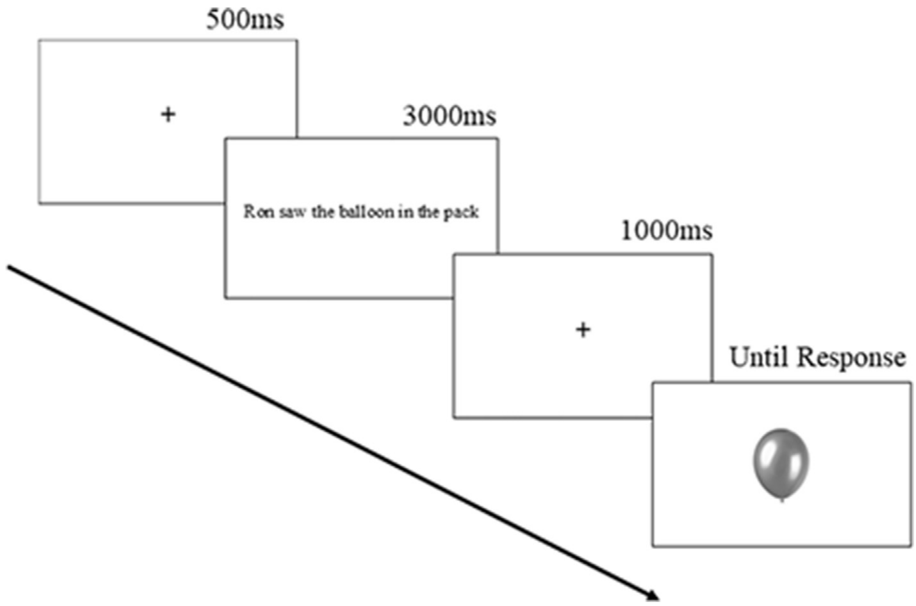


Fig. 1. The sequence of events in a trial.

[1] This long stimulus onset asynchrony (SOA) was used in order to allow sufficient time for processing the sentence. In particular, given evidence that low span readers activate a perceptual representation (of the described situation) more slowly than high span readers (Madden & Zwaan, 2006), we wanted to make sure that the participants would have enough time to activate the appropriate perceptual representation, if indeed they are able to do so. In addition, to further facilitate the task, the written sentences were also read aloud by the experimenter.

The experiment was conducted in two separate sessions (one for each condition), with an interval of over three weeks between them. At the beginning of each session, the participants were given instructions and examples, followed by 6 practice trials. Each session consisted of 46 trials - 23 related trials ("Yes" response) and 23 fillers ("No" response) - presented in random order.

In the "no-context" session, each trial started with a central fixation marker for 500 ms. The offset of the marker was followed by an ambiguous prime presented in black letters for 500 ms. Then, a central fixation marker was presented for 500 ms, followed by the presentation of a target word presented in red letters, which remained on the screen until a response was made. Participants were asked to decide as accurately as possible whether the two words (the prime and the target) are semantically related or not, by pressing the designated "Yes" or "No" keys on the keyboard.

The "context" session was identical except that a sentential context was presented before the ambiguous prime. Thus, in this session, each trial started with a presentation of a central fixation marker for 500 ms, followed by a context sentence, which was presented in black letters for 3000 ms and was read aloud by the experimenter. Then, the ambiguous prime, which was also the last word of the sentence, was centrally presented in black letters for 500 ms. The offset of the prime was followed by a central fixation marker presented for 500 ms. Then, a target word presented in red letters replaced the central fixation marker and remained on the screen until a response was made.

3. Results

3.1. Main experiment – the sentence-picture verification task

Response times (RT) and error scores to experimental trials were analyzed using a linear mixed effects (LME) model (Baayen, Davidson, & Bates, 2008). This computation allows the testing of hypotheses while taking into account the variance due to participants and to items simultaneously. Our major hypothesis relates to the difference in response times between congruent trials, in which the shape or the orientation of the object *matched* the shape or the orientation implied by the sentence, and incongruent trials, in which the shape or the orientation of the object *did not match* the shape or the orientation implied by the sentence. Thus, the model was constructed for the analysis with the effects of Group (TD adolescents vs. ASD adolescents) and Condition (match vs. mismatch) as fixed factors, and the effects of Item, and Participant as random factors. In addition, the effect of Item Type (shape or orientation), and the effects of Nonverbal IQ (The Raven Progressive Matrices (RPM) scores, Raven et al., 1998) and Verbal Ability (Kave expressive vocabulary scores, Kavé, 2005) were taken into account as fixed effects.

RT analyses: Response times of errors were deleted (2.41%), and RTs were log transformed. To examine the RT data, four models were compared in terms of fitting the data: The first model included the fixed main effects of Group and Condition and the random effects of participants and items. The second model included the fixed main effects of Group and Condition, the interactions between them, and the random effects of participants and items. The third model included the fixed main effects of Item Type, Group and Condition, the interactions between Group and Condition, and the random effects of participants and items. The fourth model included the fixed main effects of Non-verbal IQ, Verbal Ability, Item Type, Group and Condition, the interaction between Group and Condition, and the random effects of participants and items.

In order to exclude the possibility that the interaction between Group and Condition is mediated by Item Type (shape or orientation), an additional model- Model 3a - was constructed which was identical to the third model except that it included a test of the interactions between Item Type, Group, and Condition. A comparison of models 3 and 3a revealed that model 3a did not fit the data significantly better than model 3 ($\chi^2(3) = 2.25$, NS). Similarly, in order to exclude the possibility that the interaction between Group and Condition is mediated by Non-verbal IQ and/or Verbal Ability, two additional models - Model 4a and Model 4b were constructed which were identical to the fourth model except that Model 4a included all 3 way interactions between Non-verbal IQ, Verbal ability, Group, and Condition, whereas model 4b included the 4-way interaction between Non-verbal IQ, Verbal ability, Group, and Condition. A comparison of models 4, 4a, and 4b revealed no significant advantage for model 4a over model 4 ($\chi^2(9) = 10.19$, NS); In addition, model 4b showed no significant advantage over model 4a ($\chi^2(1) = 0.19$, NS). Thus, models 1, 2, 3 and 4 were compared.

The comparison between the four models revealed that the fourth model - that included the fixed main effects of Non-verbal IQ, Verbal Ability, Item Type, Group and Condition, the interaction between Group and Condition, and the random effects of participants and items - results in the best fit for the data ($\chi^2(2) = 6.34$, $p < 0.05$), relative to the third model, and was therefore selected for further analysis. Within this model, two main effects were found: (a) a significant overall effect of Item Type [$F(1, 17.48) = 12.6$, $p < 0.01$], indicating that responses to "shape" items were significantly slower (Mean log (RT) = 6.96, SE log (RT) = 0.07, mean RT = 1055.64) than responses to "orientation" items (Mean log (RT) = 6.81, SE log (RT) = 0.08, mean RT = 911.33); and (b) a significant overall effect of Non-verbal IQ (Raven scores) [$F(1, 28) = 5.96$, $p < 0.05$], indicating a negative correlation between Raven scores and RTs (the higher the score the faster the response). Importantly, the two-way interaction between Group and Condition was significant [$F(1, 1191.55) = 4.73$, $p < 0.05$], even after differences in cognitive ability were taken into account. The effect of Condition (match vs. mismatch) for each group (ASD vs. TD) is illustrated in Fig. 2. Each value represents log RT computed over subjects and items simultaneously.

In order to follow up the 2-way interaction, we analyzed the effect of Condition (match vs. mismatch) in each Group (ASD vs. TD) using the Bonferroni adjustment. As can be seen in Fig. 2, this difference was only significant in the case of TD adolescents [$\chi^2(1) = 6.07$, $p < 0.05$], as visually congruent pictures (match condition) resulted in significantly faster RTs (Mean

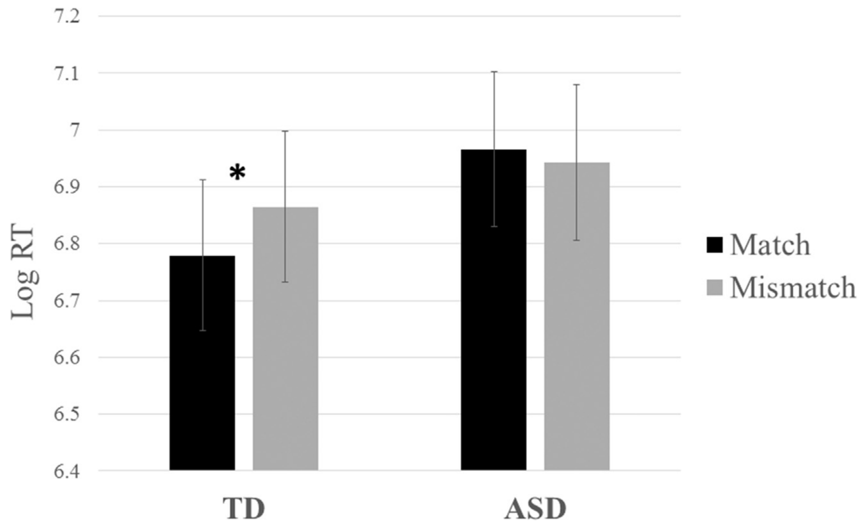


Fig. 2. Log RTs on correct responses as a function of Group (ASD vs. TD) and Condition (match vs. mismatch). Error bars indicate standard errors. *Significant, $p < 0.05$.

log (RT) = 6.78, SE log (RT) = 0.13, mean RT = 879.19) than visually incongruent pictures (Mean log (RT) = 6.87, SE log (RT) = 0.13, mean RT = 958.05). ASD adolescents, on the other hand, were not sensitive to the visual manipulation [χ^2 (1) = 0.41, NS], as congruent (Mean log (RT) = 6.97, SE log (RT) = 0.13, mean RT = 1060.19) and incongruent pictures (Mean log (RT) = 6.94, SE log (RT) = 0.13, mean RT = 1036.18) did not differ from each other.

Error Analyses: To examine the error data, the same four models were compared. Given that neither the second (χ^2 (1) = 0.12, NS), the third (χ^2 (1) = 0.77, NS), nor the fourth model (χ^2 (2) = 3.88, NS), provided a significant improvement over the first model, the model that includes only the fixed main effects of Group and Condition and the random effects of participants and items was used for further analysis. The analysis of this model revealed a significant effect of Group [F (1, 29.4) = 6.02, $p < 0.05$], indicating that the error rates were significantly larger in the ASD (M = 0.0641, SE = 0.0141), than the TD group (M = 0.0234, SE = 0.0141). The main effect of condition was also significant [F (1, 532.99) = 4.017, $p < 0.05$], indicating that visually incongruent (mismatching) pictures resulted in greater error rates (M = 0.0561, SE = 0.013) than visually congruent (matching) pictures (M = 0.0314, SE = 0.013). Although the interaction between Group and Condition was not reliable, only the TD group distinguished between the two conditions, as error rates were significantly larger than zero in the case of visually incongruent pictures, but not in the case of visually congruent pictures. ASD adolescents, on the other hand, were less sensitive to the visual manipulation, as error rates in the match condition did not significantly differ from those in the mismatch condition.

Additional analysis with matched sub-groups: Although differences in cognitive ability were taken into account in the LME analysis, to further control for this potential confound, we ran an additional analysis with matched subgroups. In particular, given that task performance was correlated with scores on the Raven test, but not with scores on the vocabulary test, we analyzed a subset of 8 ASD and 8 TD participants matched for non-verbal IQ (with scores > 90). Critically, the results with the matched subsample were consistent with our initial findings with the full sample, in that a visual match effect was found in the TD group but not in the ASD group. Specifically, in the RT analysis, even with this small sample, the two-way interaction between Group (ASD vs. TD) and Condition (Match vs. Mismatch) was (marginally) significant [F (1, 14) = 4.40, p = 0.055]. Further analysis revealed that the visual match effect was (marginally) significant in the case of TD adolescents [t (7) = 2.22, p = 0.06], as visually congruent pictures (Match condition) resulted in faster RTs (Mean log (RT) = 6.87 SE log (RT) = 0.09, mean RT = 1046.20) than visually incongruent pictures (Mean log (RT) = 7.00, SE log (RT) = 0.11, mean RT = 1275.51). ASD adolescents, on the other hand, were not sensitive to the visual manipulation [t (7) = 0.57, NS], as congruent (Mean log (RT) = 6.95, SE log (RT) = 0.16, mean RT = 1236.59) and incongruent pictures (Mean log (RT) = 6.92, SE log (RT) = 0.15, mean RT = 1210.44) did not differ from each other. The error analysis revealed no significant effects.

3.2. Additional experiment for ASD participants – the semantic judgment task

The "no context" (bank-river) and the "context" (The fisherman sat on the bank – river) conditions were compared in terms of accuracy data (for "yes" responses) across subjects (t_1) and items (t_2). Importantly, the effect of context was significant in both analyses, t_1 (15) = 5.80, $p < 0.001$; t_2 (22) = 6.51, $p < 0.001$, indicating that participants were significantly more accurate in the "context" condition (M = 0.80; SE = 0.04) then in the "no-context" condition (M = 0.59; SE = 0.04). These results suggest that participants were influenced by the biased context which facilitated their decision in the semantic judgment task. That is, the participants were able to use context in order to enhance the activation of the contextually relevant meaning.

4. Discussion

To clarify the role of perceptual simulations in language processing and to identify a potential cause of comprehension difficulties in ASD, the goal of the current study was to investigate the extent to which individuals with ASD spontaneously activate perceptual information during language comprehension. Specifically, we utilized the sentence-picture verification task (e.g., Stanfield & Zwaan, 2001) to examine the ability of adolescents with and without ASD to activate and integrate visual information about the shape or the orientation of verbally described objects, even when this information is neither explicitly mentioned nor necessary to perform the task.

Results showed that while TD adolescents responded faster when the shape or the orientation of the pictured object matched the one implied by the preceding sentence, ASD adolescents did not distinguish between the "match" and the "mismatch" conditions. Thus, consistent with previous studies (e.g., Engelen et al., 2011; Stanfield & Zwaan, 2001; Zwaan at al., 2002), we show that perceptual simulations during sentence comprehension are automatic in TD individuals. That is, TD individuals form a perceptual representation of an object's specific shape or orientation, even though making a judgment about whether or not the pictured object is mentioned in the prior sentence does not require the activation of such perceptual details. In contrast, ASD adolescents are less able to spontaneously activate contextually appropriate perceptual information in response to verbally described events.

Given that the two samples were matched for age and gender but not for general cognitive ability, the scores on both the verbal and the non-verbal tests were taken into account in the LME analysis. Importantly, the two-way interaction between group (ASD vs. TD) and condition (match vs. mismatch) was significant even after differences in cognitive ability were taken into account in the LME equation. Specifically, the LME analysis revealed a significant overall effect of non-verbal IQ (Raven scores), but no main effect of verbal ability (Kave vocabulary scores). Moreover, although non-verbal IQ was found to be correlated with RT (the higher the score the faster the response), it did not mediate the critical 2-way interaction, as the models that included 3-way and 4-way interactions between Non-verbal IQ, Verbal ability, Group, and Condition, revealed no significant advantage over the model that included only the 2-way interaction between Group and Condition (with the fixed main effects of Non-verbal IQ, Verbal Ability, Group and Condition). Thus, the LME analysis suggests that group differences between ASD and TD in perceptual simulation are unlikely to be explained by a general deficit in cognitive ability.

Although differences in cognitive ability were taken into account in the LME analysis, to further overcome the confounding effect of general cognitive ability, we ran an additional analysis with matched subgroups. Specifically, given that task performance was correlated with scores on the Raven test (non-verbal IQ), but not with scores on the vocabulary test, we analyzed a subset of ASD and TD participants matched for non-verbal IQ (with scores >90). Critically, the results with the matched subsample were consistent with our initial findings with the full sample, in that a visual mismatch effect was found in the TD group but not in the ASD group. Taken together, these two analyses suggest that the simulation deficit we observed in adolescents with ASD is unlikely to be explained by their overall lower scores on both the verbal and the non-verbal ability tests.

In addition, to investigate whether the deficit observed in adolescents with ASD is restricted to the representation of perceptual information (simulation) or includes the representation of any contextual information, we conducted a separate experiment in which we tested the effects of context on representations in a general way. In this experiment, the same ASD participants were asked to perform a semantic judgment task in which they judged whether or not a target is related to its preceding prime. The critical stimuli were related word pairs comprising an ambiguous prime (e.g., bank) and a target related to its subordinate meaning (e.g., river). Each prime-target pair was presented in two conditions: a "no-context" condition and a "context" condition. In the "no-context" condition, the prime was presented without context (e.g., bank - river). In the "context" condition, the same prime was presented after a subordinate biasing context (e.g., The fisherman sat on the - bank - river). Participants performed the task once without context and once with context. In both cases the expected answer is "yes". However, if participants are able to use context in order to enhance the activation of the contextually relevant meaning, then responses should be more accurate in the "context" condition than in the "no context" condition. This is exactly what we found. Thus, the results of this additional experiment suggest that the simulation deficit seen in individuals with ASD cannot be reduced to a general deficit in sentence-level semantic integration.

ASD adolescents' impaired ability to mentally simulate the situation described by the linguistic input can be explained within the framework of the disrupted brain connectivity theory of autism (e.g., Just et al., 2004; Just et al., 2012; Kana et al., 2006). In particular, reduced connectivity between linguistic and perceptual mechanisms may result in weaker connections between linguistic and perceptual representations. Thus, even when integration is possible at the linguistic-conceptual level, as illustrated in our additional experiment, the ability to activate perceptual representations in response to verbally described input may be impaired.

The limitation of the present study is related to the fact that our ASD group was heterogeneous in terms of cognitive ability, and overall, on both the verbal and the non-verbal tests, the mean score for the ASD group was significantly lower than the score for the TD group. Although these differences in ability were taken into account in the analysis of the results, and although there is evidence that activation of perceptual features associated with verbally described objects, in TD individuals, may be impermeable to both verbal (e.g., Engelen et al., 2011) and non-verbal abilities (e.g., Stanfield & Zwaan, 2001), still more studies are needed, with more homogeneous groups of ASD individuals (both low-functioning and high-functioning), to better address the question of whether differences in the ability to mentally simulate described situations is truly specific to ASD or is more reflective of overall cognitive function.

To summarize, the present study presents preliminary evidence that individuals with ASD may be impaired in their ability to mentally simulate verbally described situations. These findings converge with findings from other domains (e.g., social cognition) indicating embodied simulation deficiencies in individuals with ASD (e.g., Eigsti, 2013; Gallese, 2006). Clearly, more research is needed in order to evaluate the extent to which deficient simulation processes predict comprehension difficulties in ASD, and to develop effective treatment.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617–645.

Brock, J., Brown, C., Boucher, J., & Rippon, G. (2002). The temporal binding deficit hypothesis of autism. *Development and Psychopathology, 142*, 209–224.

Brosnan, M. J., Scott, F. J., Fox, S., & Pye, J. (2004). Gestalt processing in autism: Failure to process perceptual relationships and the implications for contextual understanding. *Journal of Child Psychology and Psychiatry, 45*(3), 459–469.

Dijkstra, K., Yaxley, R. H., Madden, C. J., & Zwaan, R. A. (2004). The role of age and perceptual symbols in language comprehension. *Psychology and Aging, 19*(2), 352–356.

Eigsti, I.-M. (2013). A review of embodiment in autism spectrum disorders. *Frontiers in Psychology, 4*, 224.

Engelen, J. A., Bouwmeester, S., de Bruin, A. B., & Zwaan, R. A. (2011). Perceptual simulation in developing language comprehension. *Journal of Experimental Child Psychology, 110*(4), 659–675.

Frith, U. (1989). *Autism: Explaining the enigma*. Oxford: Blackwell.

Frith, U., & Snowling, M. (1983). Reading for meaning and reading for sound in autistic and dyslexic children. *Journal of Developmental Psychology, 1*, 329–342.

Gallese, V. (2006). Intentional attunement: A neurophysiological perspective on social cognition and its disruption in autism. *Brain Research, 1079*(1), 15–24.

Giora, R., Gazal, O., Goldstein, I., Fein, O., & Stringaris, K. A. (2012). Salience and context: Interpretation of metaphorical and literal language by young adults diagnosed with Asperger's syndrome. *Metaphor and Symbol, 27*, 22–54.

Hadad, B.-S., & Ziv, Y. (2015). Strong bias towards analytic perception in ASD does not necessarily come at the price of impaired integration skills. *Journal of Autism and Developmental Disorders, 45*(6), 1499–1512.

Hahn, N., Snedeker, J., & Rabagliati, H. (2015). Rapid linguistic ambiguity resolution in young children with autism spectrum disorder: Eye tracking evidence for the limits of weak central coherence. *Autism Research, 8*(6), 717–726.

Happe, F. (1996). Studying weak central coherence at low levels: Children with autism do not succumb to visual illusions. *Journal of Child Psychology & Psychiatry, 37*, 873–877.

Happe, F. (1997). Central coherence and theory of mind: Reading homographs in context. *British Journal of Developmental Psychology, 15*, 1–12.

Happé, F., & Frith, U. (2006). The weak coherence account: Detail focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 36*(1), 5–25.

Jolliffe, T., & Baron-Cohen, S. (1999). A test of central coherence theory: linguistic processing in high-functioning adults with autism or Asperger syndrome: Is local coherence impaired? *Cognition, 71*, 149–185.

Just, M. A., Cherkassky, V. L., Keller, T. A., & Minshew, N. J. (2004). Cortical activation and synchronization during sentence comprehension in high-functioning autism: Evidence of underconnectivity. *Brain, 127*, 1811–1821.

Just, M. A., Keller, T. A., Malave, V. L., Kana, R. K., & Varma, S. (2012). Autism as a neural systems disorder. A theory of frontal-posterior underconnectivity. *Neuroscience and Biobehavioral Reviews, 36*(4), 1292–1313.

Kana, R. K., Keller, T. A., Cherkassky, V. L., Minshew, N. J., & Just, M. A. (2006). Sentence comprehension in autism: Thinking in pictures with decreased functional connectivity. *Brain, 129*, 2484–2493.

Kavé, G. (2005). Standardization and norms for a Hebrew naming test. *Brain and Language, 92*(2), 204–211.

Madden, C. J., & Zwaan, R. A. (2006). Perceptual representation as a mechanism of lexical ambiguity resolution: An investigation of span and processing time. *Journal of experimental Psychology: Learning Memory and Cognition, 32*, 1291–1303.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris, 102*(1), 59–70.

Mizuno, A., Villalobos, M. E., Davies, M. M., Dahl, B. C., & Muller, R. A. (2006). Partially enhanced thalamocortical functional connectivity in autism. *Brain Research, 1104*, 160–174.

Norbury, C. F., & Bishop, D. V. M. (2002). Inferential processing and story recall in children with communication problems: A comparison of specific language impairment, pragmatic language impairment, and high functioning autism. *International Journal of Language and Communication Disorders., 37*, 227–251.

O'Riordan, M. A., Plaisted, K. C., Driver, J., & Baron-Cohen, S. (2001). Superior visual search in autism. *Journal of Experimental Psychology: Human Perception and Performance, 27*(3), 719–735.

Peleg, O., & Eviatar, Z. (2008). Hemispheric sensitivities to lexical and contextual constraints: Evidence from ambiguity resolution. *Brain and Language, 105*, 71–82.

Plaisted, K., O'Riordan, M., & Baron-Cohen, S. (1998). Enhanced discrimination of novel, highly similar stimuli by adults with autism during a perceptual learning task. *Journal of Child Psychology and Psychiatry, 39*(5), 765–775.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's progressive Matrices and vocabulary scales*. San Antonio: Harcourt Assessment. section 1: General overview.

Snowling, M. J., & Frith, U. (1986). Comprehension in "hyperlexic" readers. *Journal of Experimental Child Psychology, 42*, 392–415.

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science, 12*(2), 153–156.

Zwaan, R. A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Moving words: Dynamic representations in language comprehension. *Cognitive Science, 28*(4), 611–619.

Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: Six replication attempts. *PloS One, 7*(12), e51382.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13*(2), 168–171.