Rapid linguistic ambiguity resolution in young children with autism spectrum disorder: Eye tracking evidence for the limits of weak central coherence

Noemi Hahn[1], Jesse Snedeker[1] and Hugh Rabagliati[1,2]

1. Department of Psychology, Harvard University, Cambridge, MA, USA

2. Department of Psychology, University of Edinburgh, Edinburgh, United Kingdom

Please address correspondence to:

Hugh Rabagliati

Department of Psychology

University of Edinburgh

Edinburgh

United Kingdom

hugh.rabagliati@ed.ac.uk

Abstract

The weak central coherence hypothesis proposes that individuals with autism have a difficulty integrating information into its broader context that is domain general. The best evidence for this difficulty's domain generality comes from reports of impaired use of linguistic context to resolve ambiguous words. However, recent work has suggested that impaired use of linguistic context may not be characteristic of ASD, and is instead better explained by co-occurring language impairments. Here we provide a strong test of these claims, using the visual world eye tracking paradigm to examine the online mechanisms by which children with autism resolve linguistic ambiguity. To address concerns about both language impairments and compensatory strategies, we used a sample whose verbal skills were strong and whose average age (7;6) was much lower than previous work on ambiguity resolution. Participants (40 with autism and 40 controls) heard sentences with ambiguous words in contexts that either strongly supported one reading or were consistent with both (*John fed/saw the bat*). We measured activation of the unintended meaning through implicit semantic priming of an associate (looks to a depicted baseball glove). Contrary to the predictions of weak central coherence, children with ASD, like controls, quickly used context to resolve ambiguity, selecting appropriate meanings within a second. We discuss how these results constrain the generality of weak central coherence.

Keywords: autism, language, lexical ambiguity, homophones, eye tracking, weak central coherence.

Rapid linguistic ambiguity resolution in young children with autism: Eye tracking evidence for the limits of weak central coherence

Much of mental life involves trying to understand things in their broader contexts. Basic perceptual tasks, like recognizing an object, can be facilitated by integrating the surrounding environment (Biederman, 1972). Social judgments, like recognition of emotion in a face, are strongly influenced by the situational context (Carroll & Russell, 1996). Aspects of language processing, such as determining the meaning of an ambiguous word like *wind,* are highly dependent on the context provided by a sentence (Swinney, 1979).

The Weak Central Coherence theory of autism (Frith, 1989; Happé, 1999; Happé & Frith, 2006) proposes that individuals with autism spectrum disorders (ASD) have a cognitive style in which processing focuses on specific details, rather than on the synthesis of information with its broader global context. As Happé (1999) puts it, in ASD "…features are perceived and retained at the expense of global configuration and contextualized meaning." Critical support for this proposal has come from studying patterns of strength and weakness in either visuospatial or auditory processing. Detail-focused processing is evidenced by: ASD participants' superior accuracy in judging the pitch of a tone (Bonnel et al., 2003); their increased ability to pick out embedded figures from a larger drawing (Jolliffe & Baron-Cohen, 1997; Ropar & Mitchell, 2001; Shah & Frith, 1983); and their faster reaction times in visual search tasks (Plaisted, O'Riordan, & Baron-

Cohen, 1998). Meanwhile, support for insensitivity to global context comes from ASD participants' higher thresholds for perceiving coherent motion in patterns of dots (Bertone, Mottron, Jelenic, & Faubert, 2003; Milne et al., 2002; Spencer et al., 2000) and their reduced use of gestalt grouping principles (Bolte & Poustka, 2006; Brosnan, Scott, Fox, & Pye, 2004; Shah & Frith, 1983).

Although weak central coherence predicts a bias in cognitive style that is domain general, there have been surprisingly few tests of the theory in domains outside of visuospatial processing and audition. The critical exception comes from research into language. To the extent that individuals with ASD show both detail focus and global insensitivity in understanding sentences, then the domain generality of weak central coherence is supported.

Weak central coherence and language

Much of the work on the use of linguistic context in ASD focuses on the resolution of ambiguous words. For instance, the word *tear* is an orthographically ambiguous homograph. Its most common pronunciation denotes a drop of liquid, as in *He shed a single tear*, but it can also denote a rip, as in *There was a big tear in her dress.* In an influential paper, Happé (1997) claimed that individuals with ASD have difficulty using context to understand ambiguous words (see also Frith & Snowling 1983). Adolescents and adults with autism (aged 8 through 28 years), along with matched controls, read aloud sentences containing homographs such as *tear*. ASD participants showed little evidence that they used

context: They rarely varied their pronunciation based on the surrounding words. In contrast, control participants offered different pronunciations for different contexts. This finding has been both replicated and extended. For example, Joliffe and Baron-Cohen (1999, 2000; see also Lopez & Leekam, 2003) found a similar reduction in the use of context in high-functioning adults with ASD. Homograph resolution has therefore been held up as evidence that weak central coherence characterizes language in autism, and is therefore a domain-general cognitive style.

Still, other researchers have cast doubt on this conclusion. Some have suggested that the homograph task may not accurately characterize language in autism (see e.g., Brock and Bzishvili, 2013; Brock and Caruana, in press). For instance, Happé's original study (and the subsequent replications) used only five ambiguous words as stimuli, and it is not clear if the results generalize outside this set. In addition, reading aloud is an indirect measure of language comprehension. It is therefore important to additionally assess the understanding of language (e.g., during silent reading or listening). Finally, the design of the study, in which both pronunciations of *tear* were used within a participant, meant that perseveration on the initial pronunciation could mask an ability to use context. In fact, Hala, Pexman and Glenwright (2007) found that children with ASD (mean age 10;4) only had difficulty pronouncing ambiguous words when they had had to switch from one pronunciation to another.

An important additional concern relates to the populations tested, rather than the task: These homograph studies did not include a detailed characterization of the linguistic abilities of the ASD groups. Based on this, Norbury (2005) proposed that difficulties resolving homographs were actually the result of limits to participants' structural language skills (the ability to combine words and phrases) rather than weak central coherence. She supported this claim with a cross-modal semantic priming study, which contrasted four groups of participants, with either an ASD diagnosis, a language impairment, neither, or both. Children (aged 9-17 years, mean 13) heard sentences containing homophones (e.g., *bank*) and then judged whether a picture (e.g., a river), presented after a 1000ms interval, fit the meaning of the sentence. Children's accuracy on this measure varied based on their linguistic ability, but not on their ASD diagnosis, consistent with the proposal that structural language skills, not ASD, determine competence at lexical ambiguity resolution.

While Norbury's hypothesis is intriguing, not all of the evidence supports it. Not only does it stand in contrast to Joliffe and Baron-Cohen's (1999) previously mentioned finding of contextual insensitivity in high-functioning individuals with ASD, but it is also inconsistent with the results of a second cross-modal priming study. Henderson, Clarke and Snowling (2011) found that a set of highly-verbal children and teenagers with ASD (aged 7 to 15, mean 11;6) failed to use context to suppress inappropriate meanings of ambiguous words. In particular, they found that ASD participants showed facilitation at naming pictures that were

related to the inappropriate meaning of an ambiguous word, while matched controls showed a delay. For, example, ASD participants were faster to name a depicted lamp after *Richard planted the bulb* than after *Chris planted the seed,* which can only be explained by the lingering presence of the unselected "light bulb" meaning. This pattern suggests that participants with ASD did not use context to select which meaning was appropriate.

The discrepancy between Norbury's and Henderson's findings may have a developmental explanation. Previous work has shown that children with ASD often learn to compensate for their early difficulties. For instance, they typically master theory of mind tasks a few years after verbal-age matched peers (Happé, 1995). Successful resolution of auditory lexical ambiguities could be another example of this: the children in Norbury's sample (who were on average two years older than Henderson's sample) might have learned to compensate for a deficit integrating spoken linguistic context. However this possibility is hard to evaluate, as the design of Norbury's study – in which priming was only assessed 1000ms after the critical word[1] – does not reveal the precise mechanisms and strategies that the groups used to integrate context.

As such, while the results of Hala et al. (2007) and Norbury (2005) throw doubt on Happé's conclusions, they do not settle the issue. We built on their findings to provide a more stringent test of the weak central coherence predictions for lexical ambiguity resolution. Like Hala and colleagues, we

attempted to minimize the effects of perseveration, this time by using a Latin square design in which each ambiguous word was only heard once. Following Norbury (2005), we evaluated lexical ambiguity resolution during auditory sentence processing, in children with strong structural language skills. However, we tried to alleviate concerns about compensatory strategies by developing a task that: a) was sensitive to the time course of lexical ambiguity resolution and b) allowed us to test considerably younger children than in prior work.

Our method was based on the visual world paradigm, in which participants' visual exploration of a scene is monitored while they hear sentences (Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995). Gaze provides a sensitive measure of how linguistic interpretation proceeds online, but it does not require participants to read or provide time-locked responses, which makes it ideal for studying children and adolescents with ASD (see also Brock, Norbury, Einav, & Nation, 2008). In order to minimize strategic effects, we measured saccades that arise as automatic responses to the linguistic input, rather than measuring eye-movements that accompany responses to spoken instructions. In particular, we built on demonstrations of implicit semantic priming in eye movements, such as the finding that participants will gaze towards pictures that are semantically or visually similar to the meanings of names that they head (e.g., looking at a key when hearing the word *lock,* or gazing at a lollipop when hearing the word *flower*) (Chen & Boland, 2008; Huang & Snedeker, 2011; Huettig & Altmann, 2004; Yee & Sedivy, 2006).

In our study, children listened to sentences while freely viewing a set of pictures. As a cover task, the last word of each sentence was missing (e.g., *Karl saw the star while he was looking for a new car, so he asked for an autograph in his…*) and participants had to choose which picture would best end the phrase. However, our experimental measure was orthogonal to this task. We used a two-by-two design that manipulated: 1) whether a target word, that appeared early in the sentence, was ambiguous or unambiguous (e.g., ambiguous: *star*, unambiguous: *actor*), and 2) whether the context before that target weakly or strongly selected the less frequent meaning of the ambiguous word (weak: *saw the star*, strong: *met the star*). Unambiguous words were synonyms of the less-frequent meaning of the target, so that they matched the context. After the target word, the sentence was identical across the four condition. It always disambiguated to the less-frequent meaning (but only five to six words later). Importantly, one of the pictures was semantically associated with the target word's more-frequent meaning (e.g., *sun* is associated with stars-at-night). We recorded participants' looks to this critical semantic associate after they heard the target word.

If participants can use preceding context to resolve ambiguous words, then the more-frequent meaning should be rapidly inhibited in the strong context condition, but it should remain active in the weak context condition. Implicit semantic priming should therefore be greater in the weak condition, and so

participants should spend more time gazing at the pictured associate (i.e., they should look more to a depicted sun when hearing *saw the star* than when hearing *met the star*). In contrast, if participants cannot use context, we would expect reliably longer looking times to the associate in the ambiguous (as compared to unambiguous) conditions, with no effect of context on the looking patterns.

We collected data on this task from large samples of children with ASD and well-matched controls (n=40 each). To ensure that any effects could not be explained by structural language delays, we followed Norbury (2005) and Henderson and colleagues (2011) and used a sample of highly verbal individuals. To mitigate concerns about compensatory strategies, we used a much younger sample than previous work on ambiguity resolution, with a mean age of 7;7. With this large, young sample and our simple, implicit task, we aimed to provide the clearest test yet of whether children with autism show a specific deficit resolving lexical ambiguity.

Method

Participants

Children with Autism spectrum disorders (ASD).

We included forty 6-to 9-year-old children with high-functioning autism spectrum disorders. Participants were recruited from the Simons Foundation's database (SFARI Base), which listed children who had participated in the Simons Simplex

Collection (SSC), a multi-site study of families in which one child has received an ASD diagnosis, while the parents and siblings have not (Fischbach & Lord, 2010). Participants were tested in hotel suites in 8 different cities in the USA.

We included children who: a) met the DSM-IV-TR (APA, 2000) criteria for either Autistic Disorder (n = 38) or Asperger syndrome (n=2), b) scored above 85 on the Core Language section of the Clinical Evaluation of Language Fundamentals – Fourth Edition, c) scored above 80 on the KBIT Test of nonverbal IQ, d) spoke English as their first language or primary language at home, e) were aged between 6-9 years, f) had vision/hearing that was either normal or corrected to normal. Diagnoses were provided by an experienced clinical team at the SSC site, and had been confirmed using the Autism Diagnostic Interview-Revised (ADI-R, Rutter, Le Couteur, & Lord, 2003) and the Autism Diagnostic Observation Schedule Module 3 (ADOS, Lord, Rutter, DiLavore, & Risi, 1999). The ADOS calibrated severity score (a standardized scoring of the ADOS accounting for age and linguistic ability, Gotham, Pickles & Lloyd, 2009) was above 5 for each child (mean=7.42). SFARI Base provided the dates on which 18 of the children were evaluated; testing occurred 17 months later (on average, range 8-23).

6 additional participants were tested but not included, because their score fell below 85 on the CELF-4 Core Language assessment (n=5) or because their score fell below 80 on the KBIT test of Nonverbal IQ (n=1).

Typically developing (TD) children.

40 TD children aged 6-9 years participated. They were recruited from the participant database of the Laboratory for Developmental Studies at Harvard University and tested there. We included children who scored above 80 on the KBIT, scored above 85 on CELF-4, spoke English as a first language, and had normal or corrected to normal hearing and vision. Parents reported that their children were developing typically, had no history of clinical diagnosis or special educational services, and were in the age-appropriate school grade. To ensure well-matched groups, we added an additional criterion for the final 10 TD children: Their core language skill score had to lie below 100. This subgroup therefore completed the CELF-IV test first, and participated in the remaining tasks only if their score fell between 85-100.

To screen for potential undiagnosed ASD, parents filled out the Social Communication Questionnaire - Lifetime Form, a 40-item questionnaire developed as a secondary screening tool for Autism based on the more extensive ADI-R (Rutter, Bailey, Berument, Lord, & Pickles, 2003),. We included children if they scored below the autism cutoff score of 15 (one potential participant was excluded).

Group matching.

Table 1 shows that groups were matched on gender, mean age, mean nonverbal

IQ (K-bit) and mean language score (CELF-IV); the distribution of CELF-IV and KBIT scores was also matched (Facon, Magis, & Belmont, 2011). We did not record participants' race/ethnicity or socio-economic status, but our informal observation was that the majority were non-Hispanic White children from middle to high socioeconomic status households.

Materials

Participants heard sentences made up of three components, an initial context (*Karl saw the*), a target word (*star*) and then a concluding context (*while he was looking for a new car, so he asked for an autograph in his…*). The initial context could be neutral or strongly selective. The target word could be ambiguous or unambiguous (matched to the subordinate meaning of the ambiguous word). For ambiguous target words, the relative frequency of each meaning was taken from norming data collected from CHILDES for Rabagliati, Pylkkänen and Marcus (2013). The concluding context, which did not vary across conditions, contained an initial section that was neutral between the two meanings, followed by a section that disambiguated to the subordinate meaning.

1) Strong Context/Ambiguous: *Karl met the star while he was looking for a new car, so he asked for an autograph in his…*
2) Weak Context/Ambiguous: *Karl saw the star while he was looking for a new car, so he asked for an autograph in his…*
3) Strong Context/Unambiguous: *Karl met the actor while he was looking for*

a new car, so he asked for an autograph in his…

4) Weak Context/Unambiguous: *Karl saw the actor while he was looking for a new car, so he asked for an autograph in his…*

Participants viewed a screen displaying four clipart pictures (Figure 1): one was related to the dominant meaning of the target ambiguous word (the prime picture, *sun*), one was a reasonable completion of the sentence (*notepad*), and two served as distracters (*candle, ipod*).

Procedure

Our task was embedded in a game. A boy called Billy had written some sentences as part of his homework, but his computer had accidentally deleted each sentence's last word. Participants helped Billy by listening to his sentences and selecting the best continuation from amongst the four pictures, using the touchscreen on the monitor. After each response, the correct picture was highlighted.

The experimental procedure was controlled using E-Prime (Schneider, Eschmann, & Zuccolotto, 2002). Sentences were presented over loudspeakers. Pictures were displayed on a 17" LCD monitor and eye-movements were recorded using a Tobii T60 eye-tracker sampling at 60Hz. Participants completed the study as part of a larger battery of seven psycholinguistic experiments and standardized tests; this was either the fourth or fifth task, depending on when the

CELF-IV was administered.

Analysis

We divided the screen into equally-sized quadrants, and coded samples on which participants gazed at the quadrant containing the prime picture as 1, gaze elsewhere as 0, and track-loss as missing data. Data were analyzed across a 1500ms window beginning at the offset of the target word, divided into 100ms bins. To remove baseline effects, we excluded trials on which participants were gazing at the prime picture at the offset of the target word (ASD = median 4 trials/participant [SD=1.7], TD= median 3[1.7]; for the motivation behind this analysis step see Tanenhaus, Frank, Jaeger, Salverda, & Masharov, [2008]). Next, we excluded trials with track-loss in more than 50% of samples in this window (median ASD=6[3.8], TD=7[5.7]). Then, for each trial, we calculated the mean proportion of looks to the prime in 100ms bins, starting from the offset of the target word. The proportion of looks was log-odds transformed to ensure it was suitable to analyze with a general linear model, with proportions at 1 adjusted to 0.975, and 0 to 0.025.

Our analyses focused on how looks varied across conditions, over time. Because we had no *a priori* time windows of interest, we corrected for multiple comparisons using a non-parametric permutation test, originally developed for EEG data (Maris & Oostenveld, 2007). The procedure uncovers contiguous clusters of statistically reliable effects and tests if those clusters were likely to

have occurred by chance.

For each bin we conducted a mixed-effect linear regression analysis on the log-odds of looking to the prime. Our predictors were target word ambiguity (ambiguous/unambiguous control), preceding context (strong/weak), subject population (typically-developing/highly-verbal ASD), and their full set of interactions. All predictors were contrast coded.[2] Because ambiguity and context were varied within-subjects, we also included them (and their interaction) as by-subject predictors. Our regression was therefore similar to a mixed ANOVA (grouping by subjects), but with the advantage that mixed models account for differences in the number of observations per cell (e.g., due to trial exclusions).

The permutation test followed the following procedure:

1) For each predictor (effects of target word, context and population, plus their three two-way interactions and the three-way interaction), find clusters of temporally-adjacent samples where the test statistic for each sample was greater than a predetermined threshold (defined below). Clusters can be as small as one sample.

2) For each cluster, sum the test statistics across its samples. Each of these summed totals will later be tested against an empirical distribution, in order to assess the probability of the cluster under the null hypothesis.

3) Permute the data, by randomizing the fixed variables while respecting its structure in other respects. Trial labels for within-subjects factors (here,

target word and context) are randomly shuffled within a subject, while trial labels for between-subjects factors (here, population) are shuffled between subjects.

4) Run steps 1) and 2) on the permuted data. Then, for each predictor, extract the largest summed test statistic from any identified clusters. These will be used to create empirical distributions for each predictor.

5) Repeat steps 3) and 4) 9999 times to create empirical distributions.

6) Take the clusters from the original data, and for each cluster from each fixed effect, compare it to the appropriate empirical distribution. The *p* value for each cluster is calculated as the proportion of permuted clusters with larger test statistics.

An advantage of this analysis is that the test statistic can be specified in advance. This meant that we could capture shallow, long-lasting effects by initially setting a critical *t* statistic threshold of 1.6, without increasing the chances of a false positive.

We additionally analyzed picture selection responses, using mixed-effects logistic regressions, with the same structure as above.

Results

We first examined which pictures children chose to complete each sentence. The ASD group chose the most appropriate picture on 83% (SD=20%) of trials and

the TD group on 90%(14%). This difference was not significant (z=1.4,*ns*). There was also no reliable difference in the percentage of trials on which each group chose the primed picture (ASD:6%[9%], TD:4%[7%], z=0.5, *ns*).

Next, we analyzed whether differences in eye-movements over time indicated different degrees of implicit semantic priming across conditions and populations (Figure 2). Recall that the mixed-effects regression was conducted at each timepoint, and the permutation test searched for clusters of timepoints at which one of the predictors (ambiguity, context, population, and their full set of interactions) was significantly larger than expected by chance.

We identified two statistically significant clusters, marked by the lines below the graphs in Figure 2. First, there was a main effect of target word ambiguity that lasted from 400ms to the end of our analysis window, at 1499ms (summed t statistic for cluster = 42.6, p<.001): During this time window, participants looked more to the prime following an ambiguous than following an unambiguous word, suggesting that across both groups of children the unselected dominant meaning of the ambiguous word was initially activated and that this caused implicit priming.

Second, there was an interaction between ambiguity and context that lasted from 500ms to 1499ms (summed t = 20.4, p<.001). As Figure 3 shows, this interaction was driven by a large reduction in looks to the prime in the strong

contexts for the ambiguous words, but little difference between the contexts for the unambiguous words.

The above effects were not qualified by any further interactions. In particular, we saw no evidence of a three-way interaction between target word ambiguity, contextual strength and population. In fact, we failed to find even a single time-point at which that interaction was reliable. Our data therefore suggests that both ASD and TD children can use context to resolve ambiguous words within a few hundred milliseconds.

To confirm this, we examined the two populations separately during the time window of the contextual strength by target word ambiguity interaction (500-1499ms). We used a mixed effects model (which again can be interpreted similarly to a mixed ANOVA) to analyze whether the average log odds of looking to the target in this time window varied based on target word ambiguity, contextual strength and the interaction of the two.[3] There was a reliable ambiguity by contextual strength interaction in both groups (TD: Ambiguous Targets: Mean proportion of time fixating on $target_{weak}$=.39[SD=.36], $\underline{M}_{strong}$=.26[.32]; Unambiguous Targets: $\underline{M}_{weak}$=.20[.25], $\underline{M}_{strong}$=.19[.26]; B = -0.23[SE=0.10], t = 2.23, p = .02; ASD: Ambiguous Targets: M_{weak} =.34[.34], $\underline{M}_{strong}$=.29[.32]; Unambiguous Targets: $\underline{M}_{weak}$=.13[.19], $\underline{M}_{strong}$=.21[.28], B=-0.24[0.10], t = 2.53, p = .01).

In the control group, this interaction was accompanied by a reliable effect of contextual strength in the ambiguous word condition (B = 0.46(0.17), *t*=2.6,*p*=.01) but not in the unambiguous condition (B = -0.01(0.14), *t*=0.5,*ns.*). For the ASD group, however, there were no reliable effects of strength within each type of word (Ambiguous: B = 0.20(0.11), *t*=1.1,*ns.*; Unambiguous: B = -0.31(0.16), *t*=1.9,*ns.*).

The simplest interpretation of these data is that both groups used context to appropriately interpret ambiguous words. This is not consistent with weak central coherence. But what if weak central coherence is only characteristic of some children with ASD, such as children with more severe presentation of symptoms? They might show a larger deficit. We therefore conducted a second analysis on the ASD group. For the time window 500-1000ms, we analyzed whether the average time participants spent looking at the prime picture varied across the conditions based on their ADOS score. We used a linear mixed effects regression, fully crossing Ambiguity, Context and ADOS score (centered and standardized) . We included random intercepts for subjects and items, and calculated *p*-values by model comparison. The three-way interaction was indeed reliable (B=-0.52(0.25), *t*=2.1,*p*=.04). However, the coefficient for this term indicates that children with higher ADOS scores showed *more* sensitivity to context, not less. Given that this result is not predicted under any theory, it is difficult to interpret (we hazard an explanation in the discussion), but it clearly does not support the weak central coherence hypothesis.[4]

Discussion

To understand language, we must resolve potential ambiguity using context. For instance, to resolve an ambiguous word, listeners must activate both of its possible meanings and then determine which is most appropriate using the surrounding context (Swinney, 1979). Children with ASD have been reported to have particular difficulties understanding ambiguous words, consistent with the weak central coherence hypothesis' claim of a domain-general difficulty integrating information into its broader context (Happé & Frith, 2006).

However, the interpretation of these studies has been controversial (Brock et al., 2008; Norbury, 2005). We used eye tracking during auditory language processing to understand whether young, highly verbal children with ASD are indeed impaired at using context. We found that young children with ASD process ambiguous words in a similar manner to matched controls. Using an implicit priming method, we found that both ASD and TD children can use strong context to inhibit the inappropriate meanings of ambiguous words. Our data suggest that they do this quickly, with evidence for inhibition emerging within 500ms of hearing the word. These results therefore contradict the predictions of weak central coherence theory. In particular, they go beyond previous work by showing that even young children with ASD are able to use context to resolve

linguistic ambiguity.

Our most surprising finding was that children who have more severe symptoms of autism also appear to make better use of context when resolving ambiguity. This result is not predicted by any theory. We think that there are two possible explanations. First, it could be a false positive (the effect is not particularly robust). But it could also be due to some as-yet-unknown protective factor(s) (such as surprisingly high processing speed or executive functions) that may characterize this subgroup. These children presumably find communicative interactions very challenging, and yet they nevertheless acquired strong structural language. Protective factors, such as high processing speed, may explain this surprising success and may also facilitate their impressive use of context.

Returning to our main finding, why did we uncover sensitivity to context, rather than what Happé and Frith (2006) called the “robustly found ASD-specific failure to disambiguate pronunciation/meaning”(p.14)? We think that part of the explanation lies in our choice of population: highly verbal individuals. The data follow the pattern in Norbury (2005): individuals with strong structural language have little difficulty resolving ambiguous words. We suspect that individuals with weak structural language would do poorly on this task.

However, while our findings (and Norbury’s) indicate that highly verbal

individuals *can* use context to resolve ambiguous words, there is other work that is inconsistent with that claim. In particular, Henderson et al. (2011) and Joliffe & Baron-Cohen (1999) found that highly verbal individuals had *difficulty* using context to disambiguate words. What can explain the difference? One likely possibility is that performance varies based on task. Our paradigm provided an implicit measure of semantic activation, reducing the explicit task demands that are found in other work (Henderson et al., 2011; Joliffe & Baron-Cohen, 1999, Norbury, 2005). Also, subtler decisions about the design may have played a large role. In particular, we used Latin square counterbalancing (so no participant saw an ambiguous word twice). Recall that Hala and colleagues (2007) found that contextual insensitivity only emerged when children had to switch from one pronunciation of a word to another. Our Latin square reduced the potential for this sort of perseverative responding (which may have affected other work, e.g., Henderson et al., 2011, Joliffe & Baron-Cohen, 1999).

What do our results mean for weak central coherence? We see two potential reconceptualizations. First, weak central coherence may not be a domain general phenomenon. It may apply to visual and auditory processing, but not language comprehension and production. Of course, it is also possible that the perceptual strengths and weaknesses of ASD can themselves be explained without weak central coherence. For instance, Plaisted (e.g., Plaisted, Saksida, Alcántara, & Weisblatt, 2003), has suggested that the salience of basic visual or auditory features is enhanced in autism spectrum disorders (see also Mottron,

Dawson, Soulieres, Hubert, & Burack, 2006). Teasing these apart requires a more mechanistic weak central coherence theory.

Alternatively, weak central coherence may exist, and may be domain general, but may only characterize a subgroup of people with ASD. This subgroup will have considerable detail-focus, and their difficulty attending to context will cause them to have language-learning difficulties. That is to say, weak central coherence may only characterize individuals with low language skills. Interestingly, this possibility predicts that the perceptual strengths and weaknesses seen in ASD, which motivated weak central coherence, should only be found in individuals with low linguistic abilities. Consistent with this prediction, adolescents with both ASD and language delay are more likely to show enhanced sensitivity to differences in auditory pitch (Bonnel et al., 2010; Jones et al., 2009).

In sum, our data suggest that highly verbal children with autism resolve lexical ambiguity as quickly and accurately as their TD peers. This constrains the generality of weak central coherence: Either it is not a domain general cognitive style, or it is not a characteristic of ASD generally. Our results also confirm suggestions that the widely-used homograph task does not provide an accurate measure of verbal ability in individuals with autism, and suggest that eye-tracking offers an impressively clear alternative measure.

References

American Psychiatric Association, APA. (2000). *Diagnostic and statistical manual of mental disorders. 4th ed.* Washington, DC: American Psychiatric Association.

Bertone, A., Mottron, L., Jelenic, P., & Faubert, J. (2003). Motion perception in autism: a "complex" issue. *Journal of cognitive neuroscience, 15*(2), 218-225. doi: 10.1162/089892903321208150

Biederman, I.. (1972). Perceiving real-world scenes. *Science, 177*(43), 77-80.

Bolte, S., & Poustka, F. (2006). The broader cognitive phenotype of autism in parents: how specific is the tendency for local processing and executive dysfunction? *Journal of Child Psychology and Psychiatry, 47*(6), 639-645. doi: 10.1111/j.1469-7610.2006.01603.x

Bonnel, A., McAdams, S., Smith, B., Berthiaume, C., Bertone, A., Ciocca, V., . . . Mottron, L.. (2010). Enhanced pure-tone pitch discrimination among persons with autism but not Asperger syndrome. *Neuropsychologia, 48*(9), 2465-2475.

Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., & Bonnel, A.. (2003). Enhanced pitch sensitivity in individuals with autism: a signal detection analysis. *Journal of Cognitive NEuroscience, 15*(2), 226-235.

Brock, J, Norbury, C, Einav, S, & Nation, K. (2008). Do individuals with autism process words in context? Evidence from language-mediated eye-movements. *Cognition, 108*(3), 896-904.

Brock, J., & Bzishvili, S. (2013). Deconstructing Frith and Snowling's homograph-reading task: Implications for autism spectrum disorders. *The Quarterly Journal of Experimental Psychology*, *66*(9), 1764-1773.

Brock, J., & Caruana, N. (in press). Reading for sound and reading for meaning in autism: Frith and Snowling (1983) revisited. To appear in J. Arciuli and J. Brock (Eds.) Communication in autism. Trends in Language Acquisition Research (TiLAR). Amsterdam: John Benjamin.

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.

Carroll, J. M, & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology, 70*(2), 205.

Chen, L., & Boland, J. E. (2008). Dominance and context effects on activation of alternative homophone meanings. *Memory & Cognition, 36*(7), 1306-1323.

Diehl, J. D., Friedberg, C., Paul, R., & Snedeker, J. (2013). The use of prosody during syntactic processing in children and adolescents with autism spectrum disorders. *Manuscript submitted for publication.*

Facon, B., Magis, D., & Belmont, J. M. (2011). Beyond matching on the mean in developmental disabilities research. *Res Dev Disabil, 32*(6), 2134-2147. doi: 10.1016/j.ridd.2011.07.029

Fischbach, Gerald D, & Lord, Catherine. (2010). The Simons Simplex Collection: a resource for identification of autism genetic risk factors. *Neuron, 68*(2), 192-195.

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

Frith, U., & Snowling, M. (1983). Reading for meaning and reading for sound in autistic and dyslexic children. *British Journal of Developmental Psychology*,*1*(4), 329-342.

Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*(5), 693-705.

Hala, S., Pexman, P. M., & Glenwright, M. (2007). Priming the meaning of homographs in typically developing children and children with autism. *Journal of Autism and Developmental Disorders, 37*(2), 329-340.

Happé, F. G. E. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child development, 66*(3), 843-855.

Happé, F. G. E. (1997). Central coherence and theory of mind in autism: Reading homographs in context. *British Journal of Developmental Psychology, 15*, 1-12.

Happé, F. G. E. (1999). Autism: cognitive deficit or cognitive style? *Trends in Cognitive Science, 3*(6), 216-222.

Happé, F. G. E., & 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. doi: 10.1007/s10803-005-0039-0

Henderson, LM, Clarke, PJ, & Snowling, MJ. (2011). Accessing and selecting word meaning in autism spectrum disorder. *Journal of Child Psychology and Psychiatry, 52*(9), 964-973.

Huang, Y. T., & Snedeker, J. (2011). Cascading activation across levels of representation in children's lexical processing. *Journal of Child Language, 38*(3), 644.

Huettig, F., & Altmann, G. (2004). The on-line processing of ambiguous and unambiguous words in context: Evidence from head-mounted eyetracking. In Carreiras, M. & Clifton, C. (Eds). *The online study of sentence comprehension: Eyetracking, ERPs and beyond* (pp. 187-207). New York, NY: Psychology Press.

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*(2), 149-185.

Jolliffe, T., & Baron-Cohen, S. (2000). Linguistic processing in high-functioning adults with autism or Asperger's syndrome. Is global coherence impaired? *Psychological Medicine, 30*(5), 1169-1187.

Jolliffe, T., & Baron-Cohen, Simon. (1997). Are People with Autism and Asperger Syndrome Faster Than Normal on the Embedded Figures Test? *Journal of Child Psychology and Psychiatry, 38*(5), 527-534. doi: 10.1111/j.1469-7610.1997.tb01539.x

Jones, C., Happé, F. G. E., Baird, G., Simonoff, E., Marsden, A., Tregay, J., . . . Charman, T. (2009). Auditory discrimination and auditory sensory behaviours in autism spectrum disorders. *Neuropsychologia, 47*(13), 2850-2858.

Lopez, B., & Leekam, S. R. (2003). Do children with autism fail to process information in context? *Journal of Child Psychology and Psychiatry, 44*(2), 285-300.

Lord, C., Rutter, M., DiLavore, PC, & Risi, S. (1999). *Autism diagnostic observation schedule*. Los Angeles (CA): Western Psychological Services.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177-190.

Milne, E., Swettenham, J., Hansen, P., Campbell, R., Jeffries, H., & Plaisted, K. (2002). High motion coherence thresholds in children with autism. *Journal of Child Psychology and Psychiatry, 43*(2), 255-263.

Mottron, L., Dawson, M., Soulieres, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders, 36*(1), 27-43.

Norbury, C.F. (2005). Barking up the wrong tree? Lexical ambiguity resolution in children with language impairments and autistic spectrum disorders. *Journal of Experimental Child Psychology, 90*(2), 142-171.

Plaisted, K., O'Riordan, M., & Baron-Cohen, S.. (1998). Enhanced visual search for a conjunctive target in autism: A research note. *Journal of Child Psychology and Psychiatry, 39*(5), 777-783.

Plaisted, K., Saksida, L., Alcántara, J., & Weisblatt, E. (2003). Towards an understanding of the mechanisms of weak central coherence effects: Experiments in visual configural learning and auditory perception. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 358*(1430), 375-386.

Rabagliati, H., Pylkkanen, L., & Marcus, G. F. (2013). Top-down influence in young children's linguistic ambiguity resolution. *Developmental psychology, 49*, 1076-1089. doi: 10.1037/a0026918

Ropar, D., & Mitchell, P. (2001). Susceptibility to illusions and performance on visuospatial tasks in individuals with autism. *Journal of Child Psychology and Psychiatry, 42*(4), 539-549.

Rutter, M., Bailey, A., Berument, S.K., Lord, C., & Pickles, A. (2003). *Social communication Questionnaire (SCQ)*. Los Angeles: Western Psychological Services.

Rutter, M., Le Couteur, A., & Lord, C. (2003). *Autism diagnostic interview-revised*. Los Angeles: Western Psychological Services.

Schneider, W., Eschmann, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools, Inc.

Shah, A., & Frith, U. (1983). An islet of ability in autistic children: a research note. *Journal of Child Psychology and Psychiatry, 24*(4), 613-620.

Spencer, J., O'Brien, J., Riggs, K., Braddick, O., Atkinson, J., & Wattam-Bell, J. (2000). Motion processing in autism: evidence for a dorsal stream deficiency. *Neuroreport, 11*(12), 2765-2767.

Swinney, D. A. (1979). Lexical access during sentence comprehension:(Re) consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*(6), 645-659.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632-1634.

Tanenhaus, M.K., Frank, A., Jaeger, T.F., Salverda, A.P., & Masharov, M. (2008). *The art of the state: Mixed effects regression modeling in the Visual World.* Paper presented at the 21st Annual CUNY Conference on Human Sentence Processing, Chapel Hill, NC.

Van Petten, C, & Kutas, M. (1987). Ambiguous words in context: An event-related potential analysis of the time course of meaning activation. *Journal of Memory and Language, 26*(2), 188-208.

Yee, E., & Sedivy, J. C. (2006). Eye Movements to Pictures Reveal Transient Semantic Activation During Spoken Word Recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 32*(1), 1-14.

Footnotes

1. Henderson and colleagues (2011) did include an additional 250ms ISI condition in which participants appeared to show typical use of context. However, it is hard to interpret the pattern of results, as it matches what one would expect if participants had not accessed the meanings of the homophones by 250ms and had based their responses on the previous sentence context alone.

2. A difficulty with this method is determining how to fully permute the data to respect the structure of a mixed within/between-subjects design. Simulations by Joshua Hartshorne suggest that the randomization done here does not inflate the Type I or II error rate, but we have also carried out an additional analysis (see supplementary materials) that solely tested for a between-population difference in the interaction score. This provided similar results.

3. Including random intercepts for subjects and items, and a maximal random effects structure.

4. We also assessed whether CELF raw scores modulated the interaction. They did not, which we attribute to limited variability in our sample's linguistic ability.

Table 1. Descriptive characteristics of the two groups (TD=typically developing and ASD=autism spectrum disorders).

Figure 1. Sample picture set for homophone “star”. The semantic associate is the sun.

Figure 2. Log-odds of looking to the primed associate over time, from the offset of the target word. Error bars indicate +/-1 standard error of the mean. The solid black lines indicate the regions where our analysis procedure finds a reliable effect of target word ambiguity, and a reliable target word ambiguity by contextual strength interaction.

Figure 3. Difference scores between ambiguous and unambiguous conditions, by contextual strength, in terms of log-odds of looking at the primed associate, plotted over time.