University of Reading

Elephants in pyjamas: testing the weak central coherence account of autism spectrum disorders using a syntactic disambiguation task

Article

Accepted Version

It is advisable to refer to the publisher's version if you intend to cite from the work.
Published version at: http://link.springer.com/article/10.1007%2Fs10803-015-2560-0
To link to this article DOI: http://dx.doi.org/10.1007/s10803-015-2560-0

Publisher: Springer

Abstract

According to the weak central coherence (CC) account individuals with Autism Spectrum Disorders (ASD) exhibit enhanced local processing and weak part-whole integration. CC was investigated in the verbal domain. Adolescent participants, recruited using a 2 (ASD status) by 2 (Language Impairment status) design, completed an aural forced choice comprehension task involving syntactically ambiguous sentences. Half the picture targets depicted the least plausible interpretation, resulting in longer RTs across all groups. These were assumed to reflect local processing. There was no ASD by plausibility interaction. Therefore, on this task, individuals with ASD did not exhibit enhanced local processing. Profiles were similar for language-impaired participants. In conclusion, weak CC may vary as a function of domain (visuo-spatial versus verbal) and task complexity.

The weak central coherence (CC) account (Frith, 1989) argues that individuals with Autistic Spectrum Disorder (ASD) exhibit a unique processing style characterised by a tendency to focus on local information. It could also be described as a difficulty establishing part-whole relationships. Weak CC is most often demonstrated using visuo-spatial tasks such as the embedded figures task. This involves finding small shapes within larger shapes, and strong performance is thought to demonstrate an ability to ignore the latter. Numerous studies have found that individuals with ASD are comparatively good at the embedded figures task, and tasks requiring similar abilities such as block design, and this may indicate weak CC (see Happé & Frith, 2006 for a summary). This processing style can account for a range of difficulties in ASD. For example, poor social communication may reflect an inability to use context (the whole) to interpret words or utterances (the part). Likewise, restricted interests may reflect a preoccupation with local information. A particularly promising aspect of this account is that weak CC is associated with severity of ASD symptoms (Olu-Lafe, Liederman, & Tager-Flusberg, 2014) and prognosis (Pellicano, 2010).

While weak CC is one of the more established theories of ASD it is not wholly uncontroversial. Firstly, it may not be exclusive to ASD. For example, individuals with eating disorders and students of drawing exhibit weak CC in the visual domain (Drake & Winner, 2011; C. Lopez, Tchanturia, Stahl, & Treasure, 2008). Secondly, it may vary across cognitive domains. The above studies found that weak CC in these non-ASD populations was specific to visuo-spatial tasks. There is also evidence for fractionation across domains within ASD, as scores on different types of CC task may not be correlated (B. Lopez, Leekam, & Arts, 2008). In fact, this study found a *negative* association between performance on a semantic memory and a face recognition CC task. Together these findings suggest that CC may not be a domain general process, or a necessary and sufficient causal mechanism for ASD.

Another debate concerns how we conceptualise CC; as reduced integration or an enhanced ability to focus on local detail (see Chamberlain, McManus, Riley, Rankin, & Brunswick, 2013; Happé & Frith, 2006 for a discussion). Integration involves combining local and global information. For example, to determine the pronunciation of a homograph (e.g. *bow* = bəʊ / baʊ) (Happé, 1997), we must identify the possible pronunciations of the word (local information) and use context (global information) to choose the right one. Reduced integration might be regarded as an impairment, as it is difficult to imagine a task where it might confer advantages. Conversely, weak CC could be conceptualised as an enhanced ability to focus on detail at the expense of the whole. This kind of ability could be regarded as a strength as it allows us to perform better at tasks such as the embedded figures task.

To evaluate these different conceptualisations of CC, we need to gather data from a variety of different tasks across different cognitive domains. Many visuo-spatial CC tasks such as the embedded figures task investigate enhanced local processing, though there are exceptions, such as motion coherence (Milne et al., 2002), which taps integration. In the linguistic domain, CC tasks focus exclusively on integration. For example, to choose the right pronunciation of a homograph (Happé, 1997; Jollife & Baron-Cohen, 1999) we must use sentence context. Joliffe and Baron-Cohen (1999) identified difficulties making bridging inferences, e.g. *George left his bath running / George cleared up the mess in the bathroom*, resulting in the inference *the bath had overflowed.* This also depends on integrating the sentence into a wider context, in this case a CAUSE-EFFECT sequence. Booth and Happé (2010) identified a tendency to complete sentences with frequent small scale chunks, e.g. *hunt with a knife and fork*, where a more appropriate response would be *hunt with a knife and gun*. Again, this is a test of integration, as context must be used to produce a meaningful fragment. Finally, studies have identified poor performance in the production and

comprehension of narratives (Barnes & Baron-Cohen, 2012; Nuske & Bavin, 2011). This may signal difficulties integrating individual story episodes into an overarching narrative structure.

Clearly CC tasks across different domains conceptualise CC in different ways with visuo-spatial tasks tending to focus on local processing, while verbal tasks tap integration. In addition, we cannot be certain how much these tasks are tapping domain-specific skills unrelated to CC. For example, linguistic CC tasks may actually reflect language abilities as opposed to CC more generally. There is converging evidence that the ability to use linguistic context during lexical disambiguation and retrieval is more closely related to overall language abilities than the presence or absence of ASD (Brock, Norbury, Einav, & Nation, 2008; Norbury, 2004). Moreover, impaired comprehension and production of narrative is also found in individuals with language-impairments and no ASD (Norbury & Bishop, 2002). Consequently findings of weak CC in the verbal domain may reflect relatively weak linguistic abilities in ASD, as opposed to weak CC.

To further investigate weak CC across domains a linguistic task was devised which, unlike previously employed tasks, tested enhanced local processing. It assessed the comprehension of sentences with Prepositional-Phrase (PP) attachment ambiguities, such as *I shot an elephant in my pyjamas.* Here the PP *in my pyjamas* can be interpreted in two ways. They either clothe the person doing the shooting, or the elephant. While, clearly the former interpretation is more plausible, the latter interpretation is also accessible, as demonstrated by Graucho Marx's famous punchline; *how he got in my pyjamas I don't know.* This example of PP-attachment is slightly unusual as the PP can "attach" to the subject *I*. In most cases the PP attaches to either the Verb or the Object. For example, in the sentence *I saw the man with the*

telescope, the telescope can be the instrument of *seeing* (the Verb), or something possessed by *the man* (the Object).

Numerous studies have investigated the phenomenon of PP-attachment. Generally, interpretation involving "high" attachment (where the PP modifies the Verb) is more common than interpretation involving "low" attachment (where the PP modifies the Object). Early theories accounted for this in terms of a parsing heuristic driven by the principle of parsimony which favours the structural interpretation corresponding to a syntactic tree with the fewest nodes (Frazier, 1987). An alternative account is that low-attachment is costly from a referential processing perspective as it involves complex presupposition (Crain & Steedman, 1985). For example, when interpreting *I saw [the man with the binoculars]$_{NP}$* (bracketing demonstrates low attachment), the listener must presuppose a set of men only one of whom has binoculars. For this reason, the high attachment interpretation is favoured.

A further important factor determining PP-attachment is the identity of individual lexical items. For example, verbs such as *fix* invite high attachment, given that fixing often involves some type of tool, e.g. *he fixed the toy with the screwdriver.* In addition, some PPs, e.g. *with the hat, with the stripes*, are more likely to denote possessions or attributes than instruments. In fact there is likely to be a complex interaction between the lexical content of the Verb and PP during comprehension. In addition, world knowledge is also likely to play an important role. For example, to interpret the sentence *he fixed the puppet with the strings* we must weigh up the likelihood that strings can be used to fix puppets versus the probability that some of the puppets do not have strings. This is clearly a very complex inferential process which is dependent on background knowledge and assumptions.

This study investigated comprehension of sentences involving PP attachment which were manipulated for plausibility. The participants heard a sentence containing a PP,

5

e.g. *the girl approached the butterfly on the log*, and immediately completed a binary forced-choice comprehension task using pictures. In the low plausibility condition the target picture portrayed the less plausible interpretation; in this case a high-attachment interpretation where the girl is using the log as a means of approaching the butterfly (she is on a river and using the log as a canoe). To identify the target picture the participant would need to suppress the more plausible interpretation, along with any background assumptions which led to that interpretation, e.g. knowledge that butterflies are often found on logs. Then they would need to reanalyse the sentence choosing the alternative structure. Consequently, the task assessed an ability to focus on local information (the two possible syntactic structures), at the expense of global information (e.g. background assumptions based on world knowledge). In this sense, it is analogous to widely used visuo-spatial CC tasks such as the embedded figures task, which involves focusing on local shapes while ignoring or suppressing larger patterns. It was hypothesised, that RT differences between the plausible and implausible conditions would reflect an individual's competence at local processing. The main hypothesis was that participants with ASD will exhibit a reduced effect of plausibility due to weak CC across domains.

The study formed part of a larger study investigating the overlap between ASD and Specific Language Impairment (Loucas et al., 2011), and consequently participants were recruited according to a 2 (ASD status) by 2 (Language Impairment status) design. This also enables an investigation of how linguistic processing differs as a function of overall language ability.

Participants

Participants with ASD and language impairments were selected from a cohort of individuals with Special Educational Needs assessed during the ██████████████████████████

6

Project ████████. ASD status was ascertained on the basis of ICD-10 criteria using the Autism Diagnostic Observation Schedule(ADOS: Lord et al., 2000), Autism Diagnostic-Inventory – Revised (ADI-R Lord, Rutter, & Couteur, 1994), clinical vignettes (brief case reports of the client made shortly after assessment) and teacher report. Each ADOS and ADI-R was reviewed by up to two independent experts (see Baird et al., 2006, for full details). Language Impairment was defined as a standard score of 77 (-1.5 SD) or below on the expressive and/or receptive subscales of the Clinical Evaluation of Language Fundamentals 3 UK (CELF-3 UK, Semel, Wiig, & Secord, 2000), while the IQ cut-off was a standard score of 80 (-1.3 SD) or above on either Performance IQ, or the Perceptual Organisational Index of the Wechsler Intelligence Scale for Children-III (WISC-III: Weschler, 1992). All participants in clinical groups passed a screening for hearing difficulties (<30dB). Parents of all participants were asked to discuss the study with their children and sign informed consent before their children's participation.

Given the time lag between the ████ study and the current study, on average 42 months, language and non-verbal abilities were retested using a shorter version of the previous assessments; Concepts and Directions (CD) and Recalling Sentences (RS) from the CELF, and Picture Arrangement (PA) and Block Design (BD) from the WISC. The two CELF subtests were chosen to measure expressive (RS) and receptive (CD) abilities, with the former being an especially reliable indicator of SLI (Conti-Ramsden et al., 2001). The WISC subtests were assessments of non-verbal ability with a minimal motor component.

Four further participants with SLI were recruited via contacts in schools with language units. ASD status was determined using the ADOS, Social Communication Questionnaire (SCQ) and the ADI-R (participants 1 to 3 only). Language and non-verbal abilities were assessed using the full WISC and CELF (participants 1 and 2), but due to time

constraints participants 3 and 4 were administered the short forms. Hearing difficulties were assessed via teacher report.

17 adolescents with TD matched on chronological age were recruited from a single school in South-West London. The school distributed information about the study and consent forms with pre-paid envelopes to the parents of all 14 year olds and participants were selected from those who completed a consent form. The shorter versions of the CELF-3UK and WISC III were used as screening measures. The SCQ was used as an autism screening measure, with no participant obtaining a score greater than 6. Each participant was offered a small cash sum to recompense their time and effort.

Additional assessments of phonological working memory (WM) (Children's Test of Nonword Repetition (CNRep): Gathercole & Baddeley, 1996), and verbal WM (Listening Recall subtest from the Working Memory Test Battery for Children (WMTB-C): Gathercole & Pickering, 2001) were conducted. It has been argued that the ability to process syntactically ambiguous sentences is related to WM abilities, as WM may be used to simultaneously maintain separate structural assignments or interepretations (Just & Carpenter, 1992). The use of WM tasks allowed us to investigate this possibility. The CNRep is a widely-used assessment which involves repeating 40 nonwords. To obtain error rates the number of words containing one or more errors is counted. The Listening Recall subtest is a complex span task similar to Daneman and Carpenter's (1980) original complex span task, which is designed to assess verbal WM. It involves listening to sentences and making true / false judgments. After a specified number of sentences participants must recall the final words of each.

Details of psychometrics are shown in Table 1 along with Tukey's HSD analyses comparing the groups. While there was good matching on the WISC, language comparisons

(CELF mean percentile) were not completely aligned with diagnostic boundaries, with the ASD-Language Typical group (ASD-LT) scoring significantly worse than the TD group, but significantly better than the ASD-Language Impaired (ASD-LI) and SLI groups. This reflects the relatively poor performance of three individuals in this group who scored < -1 standard deviations during the retest using the shortened CELF (mean percentile scores of 8.5, 8.5 and 10.5). One of these individuals obtained a standard score of 108 on the full CELF and hence the retest score can be regarded as anomalous. However, two individuals obtained scores of 87 and 85 on the full CELF suggesting borderline language abilities. The decision was made to keep these individuals in the ASD-LT group as the full CELF can be regarded as a more reliable assessment. Secondly, as the study focused on cognitive processing style in ASD, the uncertain language status of two individuals with ASD will have little impact on the main findings. However, this issue should be borne in mind when interpreting the data on group differences as a function of language status.

Stimuli

Sentence stimuli were designed according to a 2 (plausible versus implausible) x 2 (high-versus-low attachment) design. In order to determine bias towards high- or low-attachment the sentences were piloted on 20 adults with native or native-like English proficiency. For each sentence, participants were given high and low-attachment interpretations, and then asked to rate their plausibility. The scale ranged from 1 to 5, with participants selecting 1 if the high-attachment interpretation was much more plausible than the low-attachment interpretation, and 5 if the low-attachment interpretation was much more plausible than the high attachment interpretation. Scores were averaged to attain the "attachment-bias". Initially there were 12 stimuli in each block (total = 48), but 20 were dropped from the analysis due to reliability concerns (see below). The Supplementary Materials show all included items with their attachment-bias ratings. While a few of the high-

plausible and low-implausible sentences yielded mean scores in the centre of the scale, importantly the sentences in the comparison conditions all exhibited a strong bias towards either high or low-attachment.

All auditory stimuli were recorded by a native English speaker. Each sentence was matched with a pair of pictures. While one of the pictures showed the target interpretation the other acted as a foil. The foil picture depicted both the Subject + Verb + Object and the Object + Verb + PP segments of the sentence (see Supplementary Materials for an example). For example for the sentence *the girl approached the butterfly on the log*, the foil picture showed a girl approaching a butterfly, and separately, a butterfly on a log. Unfortunately, this made the foil picture more visually complex than the target picture, but was considered the best means of constructing the foil as it depicted all the open-class elements, but no additional elements, without corresponding to the target interpretation.

Stimuli were sorted into four pseudorandom orders to control for ordering effects. In each order there were no more than three consecutive sentences from the same experimental condition. In addition, placement of the target picture was pseudorandomised so that pictures appeared on a particular side (left versus right) on no more than three consecutive occasions.

Procedure

The visual stimuli were presented on a laptop computer using the DMDX / DMASTR experimental software (Forster & Forster, 2003), and the auditory stimuli were played to the participants using headphones (Pro-Luxe OA 850). Warm-up items were used to demonstrate the procedure. Participants were told "Listen carefully. You will hear a sentence and then you have to choose a picture to go with the sentence using these keys." They were shown the left and right Shift button. The experimenter then said "Listen

carefully, and try to press the buttons as quickly as possible." Warm up items consisted of high-frequency noun targets in the sentence frame *there's a(n) X*, e.g. *there's a kangaroo.* Immediately after the sound file had played visual stimuli were presented with an obvious distractor, in this case a koala. After the participants responded, the computer produced a short chime followed by a 500 ms pause before playing the next sound file. The warm-up session was re-administered for participants who had obvious difficulties with the procedure. The experiment was identical to the warm-up except for a short break in the middle. A timeout was set at 5 seconds, so that the program automatically moved on to the next item. A short break was inserted half way through the task to aid concentration.

Results

Data preparation

To investigate reliability, the number of correct responses for each item across all groups was investigated. Some items exhibited high error rates, which may have resulted from the time-pressured nature of the paradigm, and the relative complexity of the pictures. Items were dropped if less than 37 out of 61 participants (61%) identified the correct picture. This corresponds to a success rate significantly above chance according to a binomial distribution (p = 0.036*). Remaining items are shown in the Supplementary Materials.

In addition, the distribution of the RTs was inspected visually. It was found that a square root transformation was the best means of approximating a normal distribution, and consequently, RTs were square root transformed.

Analysis

All incorrect responses were dropped from the analysis of RTs. Mean untransformed RTs by group and condition are shown in Figure 1, and summary data are

shown in Table 2. There is a clear effect of plausibility, with the exception of low-attachment sentences in the Control group. In all groups except the control group the Plausibility effect is greater for the low attachment sentences. RTs are slightly longer in the ASD group.

The data were averaged by Group and Condition and investigated using a four-way mixed ANOVA. Between group effects were Language Impairment status and ASD status, both with two levels. Within group effects were Attachment Location (high versus low) and Plausibility (high versus low). In averaging by Group and Condition, the unbalanced nature of the design is dealt with as cell means are calculated across items.

There was a significant effect of ASD status ($F(1, 57) = 6.82$, $p = 0.011$, partial $\eta^2 = 0.176$) reflecting longer RTs in these participants. However, Language Impairment status was non-significant ($F(1, 57) = 0.004$, $p=0.951$, partial $\eta^2 < 0.001$), and there was no ASD by Language Impairment status interaction ($F(1, 57) = 1.69$, $p=0.200$, partial $\eta^2 = 0.05$). With regard to the experimental conditions there was a significant effect of Plausibility ($F(1, 167) = 20.7$, $p < 0.001$, partial $\eta^2 = 0.110$), no significant effect of Attachment Location ($F(1, 167) = 0.173$, $p = 0.678$, partial $\eta^2 = 0.001$), but a significant Plausibility x Attachment Location interaction ($F(1, 167) = 5.57$, $p = 0.019$, partial $\eta^2 = 0.032$). Means were investigated to determine the origins of this two way interaction. The mean untransformed RT in the implausible low-attachment condition was 2,213 milliseconds (s.d. = 795) compared to an RT of 1,846 (s.d. = 478) in the plausible low-attachment condition (difference = 367). The mean untransformed RT in the implausible high-attachment condition was 2,064 milliseconds (s.d. = 585) compared to an RT of 1,938 (s.d. = 694) in the plausible high-attachment condition (difference = 125). Consequently, it appears that the interaction is driven by a greater effect of plausibility in the low-attachment condition.

No other two- or three-way interactions attained significance ($p < 0.5$). The only term which demonstrated a trend towards significance was the three-way interaction between Language Impairment status, Attachment Location status, and Plausibility ($F(1, 167) = 2.98$, $p = 0.086$, partial $\eta^2 = 0.018$). A glance at the figure suggests that the difference in RTs as a function of Plausibility was more sensitive to attachment location in the language-impaired groups.

Further analyses investigated the relationship between the CNRep and Listening Recall and the degree of the "surprise" effect (the difference in untransformed RTs as a function of plausibility). These may be important explanatory variables, as working memory may be closely involved in the interpretation of syntactically ambiguous utterances (Just & Carpenter, 1992). The CNRep was significantly associated with the surprise effect (Pearson's $r = 0.36$, $p = 0.005**$). However interestingly, this was strongest for the individuals with strong non-word repetition skills. There was no significant association with Listening Recall ($r = 0.03$, $p = 0.800$).

Though RTs were regarded as the primary dependent variable, proportion correct was also investigated using in identical four-way mixed ANOVA design. There was a significant effect of Plausibility ($F(1, 171) = 17.2$, $p < 0.001***$, partial $\eta^2 = 0.092$) reflecting a higher success rate for plausible sentences (80%) than implausible sentences (68%). There was also a Plausibility by Attachment Location interaction ($F(1, 171) = 7.14$, $p = 0.008$, partial $\eta^2 = 0.040$). This reflects a stronger effect of plausibility in the low attachment sentences (plausible low-attachment; 86% (s.d. = 35), implausible low-attachment; 67% (s.d. = 47), plausible high-attachment; 74% (s.d. = 44), implausible high-attachment; 70% (s.d. = 46)). Both of these significant effects are consistent with the RT data.

Discussion

The study investigated the degree to which individuals with ASD and language impairments are capable of identifying the implausible interpretation of a syntactically ambiguous sentence. The dependent variable was latency, and higher latencies for implausible conditions were assumed to reflect the cost of reanalysing sentence structure. It was hypothesised that individuals with weak CC would exhibit a reduced latency as a function of plausibility, as they are better at focusing on local information (syntactic structure) while ignoring or suppressing global information (e.g. world knowledge). There were clear effects of experimental condition suggesting a higher processing load for the implausible condition. However, there was no significant interaction between any of the group factors (ASD and language impairment status) and any of the experimental conditions, suggesting that profiles did not differ across groups. The only group related effect was slightly increased latencies in the ASD groups, an effect which has been observed elsewhere in the literature (Riches, Loucas, Baird, Charman, & Simonoff, 2012).

Overall, given the lack of an ASD by plausibility interaction, there was no evidence that individuals with ASD present with weak CC in the verbal domain, when this is conceptualised as enhanced local processing. A possible implication is that enhanced local processing abilities may vary across domains. For example, much of the evidence for enhanced local processing in ASD comes from the visuo-spatial and auditory domains (Happé & Frith, 2006), and furthermore there is evidence to suggest that such skills may be domain-specific (e.g. Drake & Winner, 2011). The current findings suggest that enhanced local processing may be less evident in the verbal domain.

A discussion of the findings should be tempered by the fact that the experimental paradigm was relatively artificial. Nonetheless, the ability to focus on form while ignoring

14

semantic / contextual information, has been regarded as a good measure of metalinguistic ability, in particular the process of "control" (Bialystok & Ryan, 1985). There is evidence that metalinguistic skills are associated with overall language abilities (Finestack & Fey, 2009), and are limited in children with SLI (Lum & Bavin, 2007). Consequently, it can be argued that though the task lacked ecological validity, it nonetheless tapped an important dimension of language use. From this perspective, it is interesting to observe that adolescents with both ASD and language difficulties exhibited good metalinguistic skills, as they were able to quickly revise their interpretation of the sentences.

While the task was assumed to assess local processing during syntactic reanalysis, there is also evidence that it tapped integration during online processing. Firstly, there would have been no plausibility effect at all if the participants had not initially committed to the more plausible interpretation on the basis of lexical / semantic information. Secondly, an interesting profile was observed, whereby the effect of plausibility was greater for the low attachment sentences. This was confirmed by a statistical analysis of the interaction between plausibility and attachment location. The finding is best explained in terms of an initial bias towards high attachment (e.g. Crain & Steedman, 1985; Frazier, 1987) such that it was more costly to switch from a high to a low-attachment interpretation. Again, this suggests that the RTs are influenced by cognitive processes during the initial parse. Adopting this perspective, it can be argued that the online processing of adolescents with ASD and language difficulties is typical, as they integrate lexical information during parsing in order to commit to the most plausible interpretation, and exhibit a bias towards high-attachment which is observed in non-ASD and language-typical individuals. Nonetheless, the effect size for plausibility was substantially larger than the effect sizes any other terms (ASD status, LI status, attachment location) or their interactions. Consequently, it can be argued that the RTs predominantly reflect enhanced local processing during reanalysis.

15

An unusual relationship was observed between non-word repetition abilities and the degree of the plausibility effect. Paradoxically, those individuals with better non-word repetition abilities exhibited a greater effect of plausibility. In other words they were either more strongly biased towards the more plausible interpretation, or they found reanalysis more costly. Note that this association was not sufficiently strong to drive a group effect. For example, one might have suspected the children without language impairment who had much stronger non-word repetition skills to exhibit a larger effect of plausibility. However, the association, when collapsing across groups was moderately-sized (Cohen, 1988) and significant. Though this effect is counterintuitive, a similar pattern has been observed in experiments investigating the online processing of syntactically ambiguous utterances where individuals with strong working memory skills, assessed by reading span, are actually slower to interpret syntactically ambiguous utterances (Just & Carpenter, 1992). This may reflect a tendency to hold both structural assignments in memory until it becomes clear which one is most likely. As this consumes memory resources, only the high span individuals are capable of doing this. Presumably maintaining two structural assignments of the same sentence is advantageous from a processing perspective because, though it slows down online processing, it circumvents the costly process of reanalysis. From the perspective of the current study, the high span individuals may be engaging in a more rigorous process of syntactic reanalysis, exploiting their greater working memory capacity.

An unusual pattern was observed in the control group, which exhibited a stronger plausibility effect for high-attachment sentences. This is the opposite of the other groups. While this pattern is difficult to account for, it is possible that the bias towards high attachment was weaker in this group. Children with language difficulties may be more dependent on low-level heuristics such as word order when comprehending sentences (Evans & MacWhinney, 1999). A high-attachment bias may be another such heuristic. In fact there

was a trend towards a significant interaction between language impairment status, plausibility and attachment location, suggesting a stronger high-attachment bias in the two language impaired groups. This provides a partial explanation for a different profile in the controls, but does not explain why effects were in the opposite direction to the other groups. It is possible that with a more highly-powered study, profiles will be more consistent across groups.

Overall, the findings suggest that evidence for weak CC in ASD may vary according to the cognitive domain investigated, one's conceptualisation of weak CC as reduced integration or enhanced local processing, and also the complexity of the processing task. For example, while the inferences involved in linking sentences, or constructing a narrative (Jollife & Baron-Cohen, 1999; Nuske & Bavin, 2011) are very complex, the processing involved in the current study is arguably a lot simpler as the participants merely had to choose between two possible interpretations. In addition, performance on some CC tasks may be trainable (Drake & Winner, 2011). Overall, given the complexity of this picture, weak CC may best be regarded as a tendency, or bias, which is moderated by a wide variety of task-specific factors (Happé & Frith, 2006).

A limitation of the experimental paradigm is that the RTs reflected a series of processes, including parsing the sentence, scanning the pictures, comparing them with the original interpretation, and conducting syntactic reanalysis where necessary. Given this long chain, it is relatively difficult to determine cognitive factors affecting RTs. As mentioned above, enhanced local processing and integration may have been confounded. In addition, latencies may not have arisen due to syntactic reanalysis, but may reflect a shallow heuristic, e.g. choosing the picture with the fewest animate entities. A possible counterargument is that the interaction between plausibility and attachment location suggests that the task assesses

genuine online psycholinguistic processes. Nonetheless, it would be desirable to create an experimental task which was able to more reliably pinpoint underlying cognitive processes.

In order to more fully investigate the extent of weak CC in ASD, future studies should investigate associations between CC tasks across cognitive domains (B. Lopez et al., 2008), and select tasks which differentiate local processing from integration. In addition, it would be informative to investigate relationships between individual CC tasks and the overall severity of ASD (Olu-Lafe et al., 2014), or specific ASD symptoms, e.g. movement stereotypies. Only such a large scale multivariate approach will enable us to determine to what extent weak CC can be regarded as a core characteristic of ASD which explains numerous aspects of the phenotype.

Compliance and ethical standards

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee (South East Multicentre Research Ethics Committee (00/01/50) and De Montfort University Research Ethics Committee) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The authors have no conflict of interest to report.

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Medlrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames - the Special Needs and Autism Project (SNAP). *The Lancet*, *368*(9531), 210 – 215.

Barnes, J. L., & Baron-Cohen, S. (2012). The big picture: storytelling ability in adults with autism spectrum conditions. *Journal of Autism and Developmental Disorders*, *42*(8), 1557–1565. http://doi.org/10.1007/s10803-011-1388-5

Bialystok, E., & Ryan, E. B. (1985). A metacognitive framework for the development of first and second language skills. *Metacognition, Cognition, and Human Performance*, *1*, 207–252.

Booth, R., & Happé, F. (2010). 'Hunting with a knife and… fork': Examining central coherence in autism, attention deficit/hyperactivity disorder, and typical development with a linguistic task. *Journal of Experimental Child Psychology*, *107*(4), 377–393.

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

Chamberlain, R., McManus, I. C., Riley, H., Rankin, Q., & Brunswick, N. (2013). Local processing enhancements associated with superior observational drawing are due to enhanced perceptual functioning, not weak central coherence. *Quarterly Journal of Experimental Psychology (2006)*, *66*(7), 1448–1466. http://doi.org/10.1080/17470218.2012.750678

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Earlbaum Associates.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological parser. *Natural Language Parsing*, 320–358.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour*, *19*, 450 – 466.

Drake, J. E., & Winner, E. (2011). Realistic Drawing Talent in Typical Adults is Associated with the Same Kind of Local Processing Bias Found in Individuals with ASD. *Journal of Autism and Developmental Disorders*, *41*(9), 1192–1201. http://doi.org/10.1007/s10803-010-1143-3

Evans, J., & MacWhinney, B. (1999). Sentence processing strategies in children with expressive and expressive-receptive specific language impairments. *International Journal of Language and Communication Disorders*, *34*(2), 117 – 134.

Finestack, L. H., & Fey, M. E. (2009). Evaluation of a Deductive Procedure to Teach Grammatical Inflections to Children With Language Impairment. *American Journal of Speech-Language Pathology*, *18*(3), 289. http://doi.org/10.1044/1058-0360(2009/08-0041)

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods Instruments and Computers*, *35*(1), 116–124.

Frazier, L. (1987). Sentence processing: A tutorial review. In *Attention and Performance XII* (pp. 559 – 585). Hillsdale, NJ: MIT Press.

Frith, U. (1989). Autism: explaining the enigma.

Gathercole, S. E., & Baddeley, A. D. (1996). *The children's test of non-word repetition*. Hove, UK: The Psychology Press.

Gathercole, S. E., & Pickering, S. (2001). *Working Memory Test Battery for Children (WMTB-C)*. Harcourt Assessment.

Happé, F. (1997). Central Coherence and Theory of Mind in Autism: 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.

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

Just, A. J., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review*, *99*(1), 122 – 149.

Lopez, B., Leekam, S. R., & Arts, G. R. . J. (2008). How central is central coherence?: Preliminary evidence on the link between conceptual and perceptual processing in children with autism. *Autism*, *12*(2), 159–171. http://doi.org/10.1177/1362361307086662

Lopez, C., Tchanturia, K., Stahl, D., & Treasure, J. (2008). Central coherence in eating disorders: a systematic review. *Psychological Medicine*, *38*(10), 1393–1404.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., … Rutter, M. (2000). The Autism Diagnostic Observation Schedule, Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. *Journal of Autism and Developmental Disorders*, *30*(3), 205–223.

Lord, C., Rutter, M., & Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*(5), 659–685.

Loucas, T., Riches, N. G., Baird, G., Pickles, A., Simonoff, E., Chandler, S., & Charman, T. (2011). Spoken word recognition in adolescents with autism spectrum disorders and specific language impairment. *Applied Psycholinguistics*, *1*(1), 1–22.

Lum, J. A. G., & Bavin, E. L. (2007). Analysis and Control in Children With SLI. *Journal of Speech, Language and Hearing Research*, *50*(6), 1618–1630.

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, and Allied Disciplines*, *43*(2), 255–263.

Norbury, C. F. (2004). Barking up the wrong tree? Lexical ambiguity resolution in children with language impairments and autistic spectrum disorders. *Journal of Experimental Child Psychology*, *90*, 142 – 171.

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*(3), 227 – 251.

Nuske, H. J., & Bavin, E. L. (2011). Narrative comprehension in 4–7-year-old children with autism: testing the Weak Central Coherence account. *International Journal of Language & Communication Disorders*, *46*(1), 108–119. http://doi.org/10.3109/13682822.2010.484847

Olu-Lafe, O., Liederman, J., & Tager-Flusberg, H. (2014). Is the Ability to Integrate Parts into Wholes Affected in Autism Spectrum Disorder? *Journal of Autism and Developmental Disorders*, 1–9. http://doi.org/10.1007/s10803-014-2120-z

Pellicano, E. (2010). Individual differences in executive function and central coherence predict developmental changes in theory of mind in autism. *Developmental Psychology*, *46*(2), 530.

Riches, N. G., Loucas, T., Baird, G., Charman, T., & Simonoff, E. (2012). Interpretation of compound nouns by adolescents with specific language impairment and autism

spectrum disorders: An investigation of phenotypic overlap. *International Journal of Speech-Language Pathology*, *14*(4), 307–317.

Semel, E., Wiig, E. H., & Secord, W. (2000). *Clinical Evaluation of Language Fundamentals - Third Edition UK (CELF-3UK)*. Harcourt Assessment.

Table 1

Group characteristics, including IQ, language and Working Memory

Figures show means and standard deviations

	ASD-LI[a] n = 19 all male	ASD-LT[b] n = 11 all male	SLI n = 14 13 male	Controls (TD) n = 17 10 male	Group differences[c] (Tukey HSD)
Age in months (years; months)	176.4 (14;8) 7.4	182.6 (15;3) 8.5	181.6 (15;2) 7.0	172.8 (14;5) 4.2	TD < ASD-LT + SLI
WISC standard score[d]	106.1 12.4	108.7 12.5	109.5 15.2	105.9 14.5	No sig. diffs
CELF standard score	73.5 5.8	89.3 10.9	70.7 5.15	98.0 8.75	TD > ASD-LT + SLI + ASD-LI ASD-LT > SLI + ASD-LI
CNRep raw score	29.3 4.0	33.8 3.4	25.1 7.2	33.4 3.4	TD > SLI ASD-LT > SLI + ASD-LI
LR raw score	12.6 4.2	14.4 2.4	12.9 4.9	16.5 3.2	TD > ASD-LI

(a) ASD-LI = ASD Language Impaired

(b) ASD-LT = ASD Language Typical

(c) TD > SLI means TD scores were significantly higher than SLI scores ($p < 0.05$) according to the Tukey Test. TD < ASD-LT + SLI means that scores in the TD group were significantly lower than scores in both the ASD-LT and SLI groups.

(d) Scaled scores for two subtests were averaged, and these were converted to a standard score. The same process was applied to the language scores

Table 2

Reaction times by Group, Attachment Location and Plausibility (untransformed)

Means and standard deviations

	High Att.		Low Att.	
	Imp.	Plaus.	Imp.	Plaus.
ASD-LI	2039.6	1993.5	2232.8	1823.4
	962.0	976.8	1040.8	727.8
ASD-LT	1996.0	1764.7	2049.6	1630.8
	1033.7	826.7	920.9	852.9
SLI	2280.5	2041.2	2389.4	2060.0
	899.2	960.2	1124.8	897.4
Controls	1976.8	1662.1	1866.6	1803.9
	920.2	847.9	835.1	950.1

Figure Caption Sheet

Figure 1

Table 1

Table 2

Figure 1

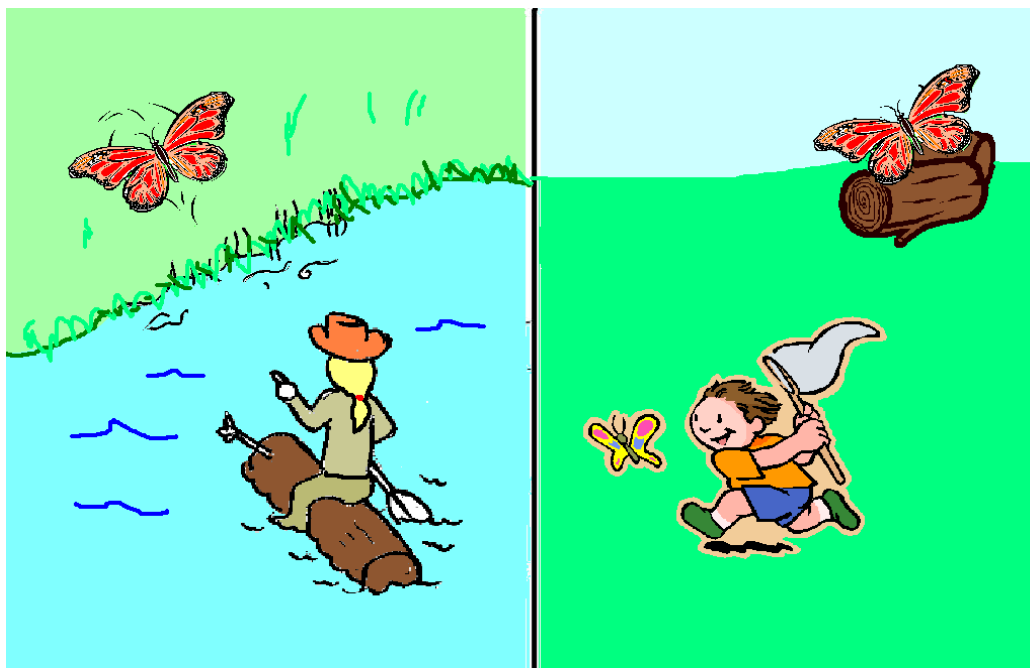Reaction Times by Group and Plausibility Condition
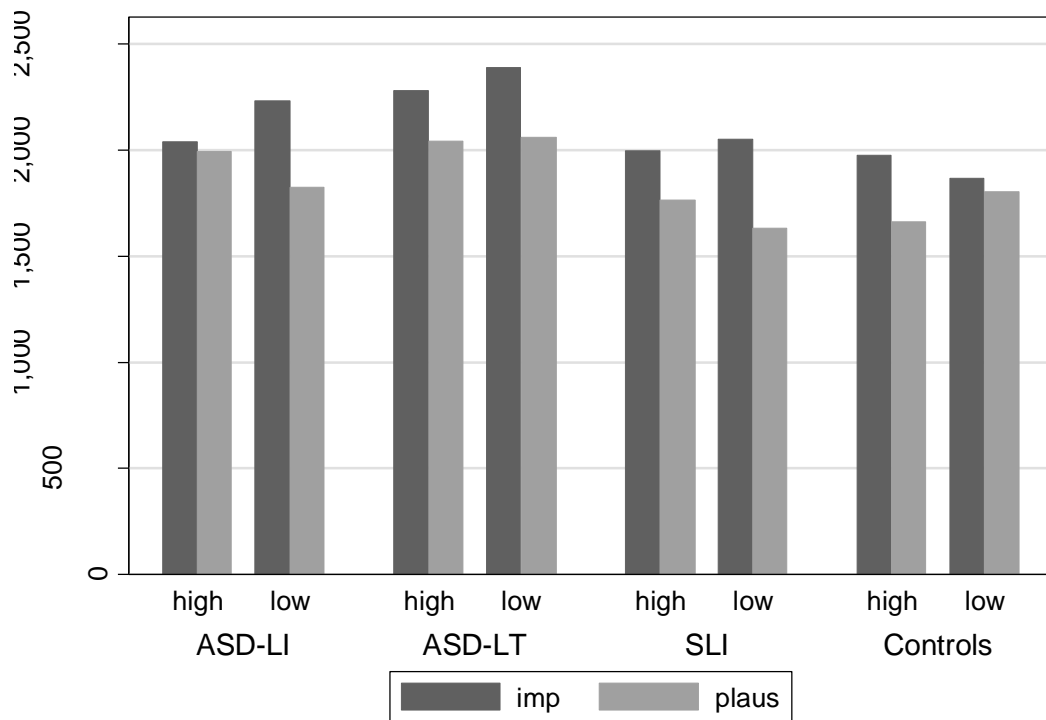
Appendix

Stimuli which passed reliability criteria (24 respondents)

High Attachment	Low Attachment
PP (underlined) modifies Verb (*in italics*)	PP (underlined) modifies NP (*in italics*)

Plausible

7. He *shot* the man <u>with the gun</u> 2.5
11. She *escaped* from the policeman <u>on the motorbike</u> 1.9
23. He *had* the sandwich <u>in the bath</u> 1.2
14. He *packed* the car <u>with his friends</u> 1.9
4. She *picked up* the frog <u>with the newspaper</u> 1.2

Mean rating = 1.74, s.d. = 0.55

Plausible

10. He pointed at *the man* <u>with the beard</u> 5.0
21. She stroked *the cat* <u>with the collar</u> 4.6
3. She saw *an aeroplane* <u>in the sky</u> 4.9
15. He pushed *the man* <u>with the glasses</u> 4.9
9. He killed *the man* <u>with the coat</u> 4.7
22. She shouted at *the kid* <u>with the cap</u> 5.0

Mean rating = 4.85, s.d. = 0.16

Implausible

6. He *attacked* the woman <u>with the hat</u> 3.6
13. She *approached* the squirrel <u>in the bush</u> 5.0
24. He *caught* the man <u>with the suitcase</u> 4.6
5. She *wrote* her name <u>on the wall</u> 4.8
18. She *ate* the spaghetti <u>with the chicken</u> 4.7
25. He *spied* the monkey <u>in the tree</u> 4.7
12. She *went* towards the butterfly <u>on the log</u> 4.8
26. He *painted* the lion <u>in the cage</u> 4.6
27. She *found* the man <u>with the dog</u> 3.9
19. He *fed* the dog <u>with the tail</u> 4.5

Mean rating = 4.52, s.d. = 0.43

Implausible

17. She climbed *the tree* <u>with the ladder</u> 1.7
8. She carried *the child* <u>on the bicycle</u> 1.6
16. He examined *the man* <u>with the magnifying glass</u> 2.3
He strangled *the man* <u>with the scarf</u> 2.2
20. She chased *the elephant* <u>on the bicycle</u> 1.7

Mean rating = 1.9, s.d. = 0.32

Example picture for stimulus *She went towards the butterfly on the log*