The present study evaluated the functioning and structuring of semantic knowledge for natural categories in Autism Spectrum Disorders (ASD). A sample of high-functioning adults (HFA) with ASD (n = 20) and a sample of typical developed individuals (n = 20) (matched for age, education, and verbal IQ) performed two tasks with natural concepts: a task of semantic access from pictures (the Camel and Cactus test) and a semantic matching-to-sample task, designed to evaluate category boundaries and structure. Overall, the HFA group showed preserved access to conceptual knowledge but some irregularities in category structure as reflected by the increased in typical errors found in the matching-to-sample task. This pattern of under generalization is suggestive of categories that have an abnormal and restricted content, missing out the typical instances.

Keywords: autism spectrum disorders, categorization, semantic memory, typicality

Semantic memory is a long-term memory system that organizes and stores our knowledge about the meaning of words, objects, facts, and people (Martin & Chao 2001; Tulving, 1972). Conceptual knowledge in semantic memory includes information about individual instances as well as the relations between them, which are a product of an extraction process of common properties (Murphy, 2004). In this context, categorization processes are especially important to promote cognitive economy, as they enable the grouping and organisation of individual items in our semantic system (Murphy, 2004; Rosch, 1978). Categorization is thus a central function of the semantic system that allows us to efficiently recall, use and produce information needed for language, communication and reasoning (Martin & Chao, 2001; Murphy, 2004; Rosch, 1978).

O’Riordan and Plaisted (2001) consider that categorization processes may be impaired in individuals with Autism Spectrum Disorders (ASD) as they showed enhanced discrimination abilities on a visual search task. In particular, higher discriminability could result in a reduction of the amount of information that can be generalized and extrapolated from one situation to another, which is detrimental to categorization (O’Riordan & Plaisted, 2001). In agreement with this...
perspective, children with ASD have been found to have difficulties in categorizing stimuli according to more abstract criteria like sports vs. games (Ropar & Peebles, 2007). The presence of categorization abnormalities supports the idea that there may be a deficit in the development and organisation of the conceptual system underlying the language and communication disturbances commonly found in ASD (Tager-Flusberg, 1996). The present study aimed to evaluate this possibility, considering both access and structure of conceptual knowledge in adult individuals with ASD.

Several studies have tested the ability of ASD individuals to learn new artificial categories and form prototypes, corresponding to a summary representation of multiple experienced items that go together in a category (Froelich et al., 2012). A mixed pattern of results has emerged from these studies with some reporting no prototype effects (e.g., Church et al., 2010; Klinger & Dawson, 2001) while others report significant prototype effects (e.g., Froehlich et al., 2012; Molesworth, Bowler, & Hampton, 2005) or preserved non-linear categorical classification curve of perceptual stimuli (Soulières, Mottron, Saumier, & Larochelle, 2007), similar to what is found in individuals with typical development.

Nevertheless, these latter studies showed some abnormalities in categorization for ASD individuals. For example, Soulières et al. (2007) did not find for this group the usual discrimination pattern of facilitation around the category boundaries observed in individuals with typical development. Froelich et al. (2012) reported difficulties in categorization of highly distorted items as compared to a control group, and suggested that although there was an intact prototype formation in the ASD group, there were signs of impaired generalization. Moreover, Molesworth, Bowler, and Hampton (2008) results question the homogeneity of ASD individuals’ performance in terms of the prototype effect. While two thirds of their ASD sample showed a standard prototype effect, the remaining group, defined by their performance on a control task, did not show the effect.

Another set of studies have reported preserved formation of new categories in individuals with ASD but consistently found that this group was slower and needed higher exposure to the materials in terms of the amount of training sessions during the learning phase (e.g., Bott, Brock, Brockdorff, Boucher, & Lamberts, 2006; Soulières, Mottron, Saumier, & Larochelle, 2011; Vladusich, Olu-Lafe, Kim, Tager-Flusberg, & Grossberg, 2010).

Taken together these studies indicate that although there are no consistent or gross deficits in the ability to categorize new material in ASD individuals, they learn slower than typically developed individuals and nevertheless show some abnormalities in category structure and organization. However, the fact that these studies have used artificial categories may not be informative of ASD individuals’ performance with natural categories. In fact, artificial categories, being usually simpler than natural categories, may have offered an easier learning context and, as such, overestimated ASD group performance.

Few studies have evaluated the structure of natural categories in ASD individuals and the findings are mixed. For example, Tager-Flusberg (1985) showed that ASD children categorize items accurately both at basic and superordinate levels. Moreover, when categorizing items that varied in typicality, a normal pattern of both under and overgeneralization was reported. In particular, children with ASD and a control group showed a similar pattern of under-generalization errors, i.e., missing to correctly identify peripheral exemplars (e.g., categorizing a penguin as a bird). Also, no differences emerged between the two groups in over-generalization errors, such as incorrectly identifying items that shared perceptual or functional features as members of the category (e.g., categorizing a whale as a fish).
In another study using natural categories, Gastgeb, Strauss, and Minshew (2006) showed that both children and adolescents with ASD were slower in classifying atypical exemplars, as compared to control participants. This latter pattern of results led Gastgeb and Strauss (2012) to suggest that in individuals with ASD, the categories had an abnormal structure and fuzzier boundaries that impaired categorization of atypical members, which fall at category boundaries. As such, it is not clear whether categorization performance of ASD individuals in natural categories is impaired in comparison to individuals with typical development and, if so, on which particular aspects or functions of semantic memory these differences emerge.

The present study aimed to evaluate two different functions of the semantic system in ASD in the context of natural categories: the access to semantic information related to a given concept (task 1), and the structuring of natural categories (task 2). For this purpose, we tested a group of high-functioning young adults with ASD and a group of typically developing individuals matched for of age, gender, education level and verbal IQ in the Camel and Cactus test (Bozest, Lambon-Ralph, Patterson, Garrard, & Hodges, 2000). In this task, participants are required to explicitly select from four alternative images which one is more semantically associated to a target image (task 1). The Camel and Cactus tests was design along the principles of the Pyramids and Palm Tree test (Adlam, Patterson, Bozest, & Hodges, 2010). This latter test can determine the degree to which an individual can access meaning from pictures (or words). Information from the Camel and Cactus test will enable us to establish whether a subject’s difficulty in naming or pointing to a picture is due to a difficulty in retrieving semantic information from pictures (Howard & Patterson, 1992).

Participants were further tested in a semantic matching-to-sample test (adapted from Lambon-Ralph, Sage, Jones, & Mayberry, 2010) that was designed to pit surface similarity against category membership (task 2). Items that vary in typicality (typical and atypical items) were presented along with items that have shared superficial similarities (pseudo-typical items). Importantly, task 2 allows us to examine the location of category boundaries by evaluating simultaneously under- and over-generalization errors. In the first case, the reduced surface similarity of atypical items would lead to a failure to identify an object as belonging to the category. In the case of pseudo-typical items, the surface similarity would lead objects to be swallowed up into a wrong category.

Consistent with earlier results showing compromised generalization (e.g., Froehlich et al., 2012) it is expected that the ASD group will show more under-generalization errors, reflecting the presence of conceptual categories with abnormal boundaries and more limited content. Critically, this difference in category boundary and content could nevertheless be independent of the amount of knowledge associated with a given category (as assessed by task 1), a separate function of the semantic system that is expected to be preserved in ASD in accordance with previous literature.
Methods

Participants

Twenty high-functioning young adults with ASD (HFA) (one female) and twenty young adults with typical development (one female) participated in the study. HFA were selected if they scored above 70 points in both the verbal and performance subscales of the Wechsler Adult Intelligence Scale (WAIS), and had been diagnosed with Autism or Autism Spectrum Disorders. The diagnosis of autism was based on DSM-IV criteria (American Psychiatric Association, 1994). The diagnostic evaluations and interviews were independently performed by two observers with large clinical experience in autism. To be included in the study, the diagnosis had to be confirmed by both observers. Asperger’s Syndrome Diagnostic Scale (ASDS, Myles, Bock, & Simpson, 2001), and/or the Autism Diagnostic Observation Schedule (ADOS, Lord, Rutter, DiLavore, & Risi, 1999) were used to confirm the clinical evaluation diagnosis.

Both groups of participants were matched on age (p > .9), verbal IQ (WAIS, p > .1), and education, measured in number of school years (p > .7) (see demographic table below).

Table 1
Demographical Information about the Two Groups. Mean (Standard Deviation) and Range are Provided for Each Group (N = 40)

<table>
<thead>
<tr>
<th></th>
<th>Controls</th>
<th>HFA</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Age (years)</td>
<td>25.05 (7.63); 18-41</td>
<td>25.25 (6.71); 18-42</td>
</tr>
<tr>
<td>Verbal IQ (WAIS)</td>
<td>109.9 (13.82); 75-128</td>
<td>103.8 (9.11); 88-126</td>
</tr>
<tr>
<td>Education (years)</td>
<td>12.9 (2.34); 9-18</td>
<td>13.15 (2.06); 10-17</td>
</tr>
</tbody>
</table>

Participants with ASD were recruited through the clinical database of the CADIN – Centro de Apoio ao Desenvolvimento Infantil (Cascais, Portugal). Control participants were recruited from the community by public advertisements and flyers and were administered an intelligence scale (WAIS, verbal subscale) by the experimenters. Ethical permission for the study was granted by the Ethical committee of the Faculty of Psychology – University of Lisbon (Portugal). Each participant gave written informed consent, where they were informed among other things of the aim of the study, that their participation was voluntary, and that they could leave any task at hand without the need for an explanation.

Materials and Procedure

Task 1: Camel & Cactus test. The Camel and Cactus test comprises 64 items (Bozeat et al., 2000). Two of the items were adapted to the Portuguese culture, as they involved information that was culture-specific. In this task, one target coloured image (displayed at the top of the screen) and four alternative coloured images are presented on each trial. Participants were asked to select which of the four same-category alternative images was more associated with the target item. For example, in one of the trials, the target “camel” must be paired with one of
four types of plants: “cactus” (the correct answer), “tree”, “sunflower” or “rose”. Alternatives were identified with the numbers one to four and participants were asked to press one of four corresponding keys. At the beginning of the test the instructions were presented on the computer screen and four practice trials were run and supervised by the experimenter. If instructions were understood correctly the task started and the 64 trials were presented in random order. Presentation and timing of stimuli were controlled using EPrime 2 software (www.psnet.com). Both accuracy and response times were registered.

**Task 2: Semantic matching-to-sample.** The task comprised 26 trials, from which 13 were living and 13 non-living target items. In each trial an array of nine numbered coloured images was displayed. The number of target items presented in each array varied between two to four to prevent any strategies in terms of number of targets per trial; two to four images were basic foils, and one to four were unrelated items (n = 54), that were from a different domain than the target item. The target items were either a typical item of a given category or an atypical item from that same category. Typical and atypical items were selected from a pre-test, in which participants saw images of exemplars and had to evaluate in a 5-point scale whether the item was a good or bad exemplar of a given category (from one – very typical to five – very atypical). The participants included in the pre-tests did not took part in the actual study. A total of 45 typical items and 36 atypical items were selected. Paired sample t tests confirmed that the typicality ratings were significantly different between the typical and atypical set (Typical items: M = 1.34, SD = .23; Atypical items: M = 3.92, SD = .75, p < .001).

Foils comprised both Pseudo-typical items (n = 57) and Partially-related items (n = 42). Pseudo-typical items were items that belonged to the same domain, had shared features with the target item, were visually similar and belonged from the same superordinate level. Partially-related items were also from the same domain and superordinate level but shared fewer features and were visually dissimilar to the target items. A second pre-test was run to evaluate the visual complexity of all images (n = 234). Participants were asked to evaluate in a 5-point Likert-scale (from one – very simple to five – very complex) the complexity of the image itself, in terms of detail, elaboration and relation between the parts (and not the object depicted in it). A t test comparisons showed that target images (M = 2.99) did not differ in terms of visual complexity from the Pseudo-typical items (M = 2.81, p = .2), Partially-related items (M = 2.90, p = .49) and Unrelated items (M = 3.00, p = .928). Before the beginning of the task, instructions were displayed on the computer screen and one practice trial was performed. Participants were told that in each trial they would have to choose (as many images as they wanted, from one to nine) the images that corresponded to a probe-name (e.g., Spiders) that was shown on the top of the screen. Participants responded by pressing the corresponding numbers’ buttons. Their responses would appear on the bottom of the screen and could be corrected. Having selected one or more items, participants were then instructed to press Enter when they were certain of their answer and to proceed to the next trial.

The two tasks were applied in the same session, and the order of administration was counterbalanced between subjects.
Results

Task 1: Camel and Cactus Test

In the Camel and Cactus test both accuracy (percentage of correct answers) and response times (in milliseconds) for correct answers were analysed (Figure 1). Although accuracy was slightly higher for the control group (M = 88.98, SD = 6.13) as compared to the HFA group (M = 86.33, SD = 11.55), independent sample t tests revealed that this difference was not statistically significant (t(1,38) = 0.91, p > .3). Incorrect answers were excluded from the response time (RT) analyses as well as answers that fell out of a 2 standard deviation from the grand mean for each participant. The RT analyses showed that it took significantly longer for the HFA group to respond (M = 8700.34, SD = 3952.27) when compared to the control group (M = 6066.13, SD = 2206.98) (t(1,38) = 2.60, p = .013).

Our findings indicate that this semantic access task is more difficult for HFA than controls. Alternatively, it may also suggest that the HFA group performed it more carefully, as revealed by larger RTs. However, the fact that both groups showed unnoticeable differences in accuracy suggests that conceptual knowledge seems to be intact and accessible in HFA individuals.

Task 2: Semantic Matching-to-Sample

For each participant we calculated the total number of errors and the proportion of different error types. Participants could either fail to select an item as belonging to a particular category or to incorrectly select an item as belonging to the category. There were five types of

---

1 For n equals 20 the normality of the data can be assumed and it is safe to use parametric tests (Myers & Hansen, 2012). Moreover, the normality of residuals was analysed with the Shapiro-Wilk test, for Total errors, and p values are higher than .001. Regarding the homogeneity of the variance the results do not differ if assuming unequal variances, and are reported accordingly.
errors: two were omission errors (i.e., participants failed to select either the typical or the atypical items as target items); three were commission errors (i.e., participants incorrectly selected as targets either pseudo-typical items, partially-related items, or unrelated items). Independent sample t tests were run for each error type given that the total number of possible error differed between categories. Errors of omission of typical items differed significantly between the groups ($t(1,38) = 2.17, p = .039$). No differences were observed between groups for either atypical items, pseudo-typical items, partially related items, or unrelated items (all $ps > .26$).

![Error bar graph](image)

**Figure 2.**
Mean of errors for the HFA and the control group for the five different possible error types. Error bars represent standard deviation from the mean.

The fact that the amount of errors differed significantly between the HFA and control groups for typical items suggests that, in accordance with our predictions, there is a clear pattern of under-generalization of concepts.

**Discussion**

In the present study we aimed at evaluating different functions of the semantic system in high-functioning young adults with ASD. In task 1 we tackled the knowledge associated with semantic concepts and in task 2 we evaluated the content of natural and existing categories and of their boundaries. A differential pattern emerged as we found that while the access to semantic knowledge was intact in high-functioning individuals with ASD, when given enough time to successfully retrieve it, the access to the content of natural categories in this population seems to be anomalous.

Results from task 2 indicated that categories were in fact irregularly formed with a significant increase of errors in typical items in individuals with ASD. This pattern is consistent with the existence of categories in which the boundaries are irregular and highly restricted.

Few studies have analysed the content of natural categories and our data is apparently in contrast with the results of Tager-Flusberg (1985), who observed normal categorization for categories of both basic and superordinate level. In this study however the trials had always a fixed number of correct responses, which could have led to the development of strategies,
leading to an underestimation of both under- and over-generalization errors. Importantly to our data, a fixed number of correct responses could have resulted in a reduced number of miss responses of the target items (both typical and atypical), which could explain in part the inconsistency with the study of Tager-Flusberg (1985).

Besides, considerable heterogeneity has been recognized as a characteristic of ASD. In particular, in tasks where developmental variables can contribute to performance such as in the tasks used here that evaluated the content of natural categories (as opposed to tasks that employ artificial categories), a less homogeneous performance can be expected.

The fact that a poor performance of high-functioning individuals with ASD in comparison with the matched control sample was only found for typical exemplars of a given category is quite puzzling. A full pattern of under generalization could be expected with substantial differences both on typical and atypical items. This was not case, and surprisingly individuals with ASD failed on the most regular exemplars. Thus, we cannot rule out the explanation that atypical items are being stored as individual items and not as a member of a category, and this hypothesis would benefit from further testing.

An alternative way to look to our results is by taking into account a recent proposal – the hyper-systemising theory (Baron-Cohen, 2006). According to it, the level of systemising that span in a continuum in the population is set high in individuals with ASD. This would provide an explanation to why individuals with ASD have preference for highly lawful systems and explain why they have trouble when faced with complex systems or less lawful change (e.g., emotions) (Baron-Cohen, 2006). The hyper-systemising theory hence argues that the excellent attention to details is directed towards extracting rules and has a positive and functional role. This account predicts that over time and experience, individuals with ASD may achieve an excellent understanding of a whole system (Baron-Cohen, 2009). In the case of the stimuli used in our experiment it could be the case that they are not considerably lawful not allowing the full extraction of regularities.

Moreover, the fact that we used coloured images (as well as Gastgeb et al., 2006) could have led to an attentional bias towards local processing and attention to details in individuals with ASD (see for instances Brown & Bebko, 2012).

The identification of variables that affect performance on this task is a limitation of the current study. Ideally a comparison to the scores obtained in autism specific scales and, in particular, to their sub-scales, could be of value and highly informative. Hence and despite the dissociation found between the two semantic functions evaluated, where the HFA group performed at comparable levels as the control group when accessing information related to a core concept, in contrast with a poorer performance when having to categorize items correctly, it would be premature to conclude that these two functions of the semantic system can be selectively preserved/impaired. Nevertheless, the study of semantic related skills in autism spectrum disorders could be a promising field providing a study model for the functioning of the semantic system.

In agreement with Gastgeb and Strauss (2012) it could be the case that individuals with ASD could gain with teaching formats that stress the formation of appropriate categories and correct category membership, with for instance, the explicit teaching of both typical exemplars and less typical members This experience could promote a broader understanding of word knowledge, and contribute in this way to the improvement of language skills in the ASD population.

One limitation of the current study is the fact that we have tested only high-functioning individuals with ASD which it was a methodological option, but prevents us from generalizing
to the whole spectrum. Conversely, by doing this so, allowed us to more reliably interpret our results and exclude the involvement of deficits (e.g., comprehension) that were not under scrutiny. Another limitation to our study is the fact that we matched our samples, only, on an IQ verbal subscale, and not on the performance IQ.

Taken together, the present study seems to point out a pattern of both preserved and impaired semantic function in high-functioning individuals with ASD. On the one hand results suggest that these individuals with ASD have intact access to conceptual knowledge associated with core concepts. On the other hand, results suggest that the access to the content of their natural categories is more restricted (as characterized by under-generalization errors). This finding could lead to the development of more explicit teaching techniques and/or promote the exposure to higher lawful systems (see also Baron-Cohen, 2009) that could overcome this apparent deficit.
References


