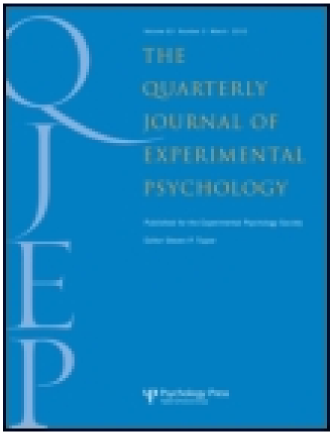

The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

Perceptual similarity in autism

Lewis Bott [a], Jon Brock [a], Noellie Brockdorff [a], Jill Boucher [a] & Koen Lamberts [a]

[a] University of Warwick, Coventry, UK

Published online: 17 Feb 2007.

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2006, 59 (7), 1237–1254

Psychology Press
Taylor & Francis Group

Perceptual similarity in autism

Lewis Bott, Jon Brock, Noellie Brockdorff, Jill Boucher, and Koen Lamberts
University of Warwick, Coventry, UK

People with autism have consistently been found to outperform controls on visuo-spatial tasks such as block design, embedded figures, and visual search tasks. Plaisted, O'Riordan, and others (Bonnel et al., 2003; O'Riordan & Plaisted, 2001; O'Riordan, Plaisted, Driver, & Baron-Cohen, 2001; Plaisted, O'Riordan, & Baron-Cohen, 1998a, 1998b) have suggested that these findings might be explained in terms of reduced perceptual similarity in autism, and that reduced perceptual similarity could also account for the difficulties that people with autism have in making generalizations to novel situations. In this study, high-functioning adults with autism and ability-matched controls performed a low-level categorization task designed to examine perceptual similarity. Results were analysed using standard statistical techniques and modelled using a quantitative model of categorization. This analysis revealed that participants with autism required reliably longer to learn the category structure than did the control group but, contrary to the predictions of the reduced perceptual similarity hypothesis, no evidence was found of more accurate performance by the participants with autism during the generalization stage. Our results suggest that when all participants are attending to the same attributes of an object in the visual domain, people with autism will not display signs of enhanced perceptual similarity.

Autism is characterized by deficits in reciprocal social behavior, communication, and behavioural flexibility (American Psychiatric Association, 1994; Wing, 1996), but is also associated with certain strengths, particularly in performance on visuo-spatial tasks such as block design (e.g., Shah & Frith, 1993; Tymchuk, Simmons, & Neafsey, 1977), the embedded figures task (Jolliffe & Baron-Cohen, 1997; Shah & Frith, 1983; but see Brian & Bryson, 1996), and visual search tasks (O'Riordan & Plaisted, 2001; O'Riordan et al., 2001; Plaisted et al., 1998b). It has been argued that studying these cognitive strengths in autism may be particularly informative because, unlike deficits, they cannot readily be explained in terms of general mental retardation (Happé, 1999b).

Frith (1989; see also Frith & Happé, 1994; Happé, 1999a) proposed that many of the cognitive strengths and weaknesses in autism could be

Correspondence should be addressed to Lewis Bott, Department of Psychology, New York University, 6 Washington Place, New York, NY 10003, USA. Email: Lewis.Bott@nyu.edu

Lewis Bott is now at the Department of Psychology, New York University. Noellie Brockdorff is now at the Centre for Communication Technology, University of Malta, Malta. Jon Brock is now at the Department of Experimental Psychology, University of Oxford. Research was supported by an RDTF grant to Lamberts, Boucher, and Bott from the University of Warwick. Lewis Bott was supported by a studentship from the Brain and Behavioral Sciences Research Council and by a grant from the Centre National de la Recherche Scientifique (France) as part of Action Thematique et Incitative. Jon Brock was supported by a studentship from the Williams Syndrome Foundation.

http://www.psypress.com/qjep
DOI:10.1080/02724980543000196

understood in terms of what she called "weak central coherence"—a tendency to focus on local details at the expense of the global "big picture" and to process information in isolation from its context. According to this account, people with autism perform well on the embedded figures task because this involves searching for a local target in a global picture. Similarly, they perform well on the block design task because they find it relatively easy to break the global target pattern down into its constituent local parts. However, weak central coherence struggles to explain enhanced visual search performance in autism. Moreover, evidence from other paradigms tends to show evidence for enhanced local processing but little evidence for impaired global processing (e.g., Mottron & Burack, 2001).

More recently, Plaisted and colleagues (O'Riordan & Plaisted, 2001; O'Riordan et al., 2001; Plaisted, 2001; Plaisted et al., 1998a, 1998b) have proposed that findings previously attributed to weak central coherence may be better understood in terms of reduced perceptual similarity.[1] Thus, detection of the target shape in the embedded figures test is relatively easy because the target is seen as being relatively dissimilar to other shapes in the overall picture (Plaisted, 2001). Similarly, superior performance on visual search tasks can be explained in terms of enhanced ability to discriminate between the target item and similar distractor items (O'Riordan & Plaisted, 2001, cf. Wolfe, Cave, & Franzel, 1989). According to this account, reduced perceptual similarity also entails a deficit in generalization such that novel objects and events are seen as highly distinctive from previous experiences. Plaisted (2001) therefore suggested that reduced similarity could explain poor generalization of social training to real-life situations (e.g., Ozonoff & Miller, 1995; Swettenham, 1996) and the restricted range of interests shown by many people with autism.

In support of this account, Plaisted et al. (1998a) reported a perceptual learning study in which, unlike controls, high-functioning adults with autism failed to benefit from preexposure to similar stimuli. The authors assumed that similarity between stimuli was an increasing function of the number of shared features and a decreasing function of the unshared features (cf. Tversky, 1977). They argued that individuals with autism failed to generalize from the preexposure phase to the test phase because they were preferentially attending to the features that discriminated between stimuli and therefore failed to notice the similarity between the sets of stimuli in the preexposure phase and the test phase of the task. The difficulty for this account is that the two stimuli in the preexposure phase and the two stimuli in the test phase all differed on the same features. Thus, by attending to the features that discriminated between stimuli in the different phases, participants with autism would also have been attending to the features that enabled discrimination between the two stimuli within the transfer phase and should therefore have shown an enhanced transfer effect. An alternative explanation for this finding is that participants with autism were simply more likely than controls to follow the experimenter's instruction to treat the test phase as a new task. This would reflect differences in strategy (or naivety to experimental manipulations) rather than perceptual abnormalities.

The aim of the current study was to formalize and then test the reduced perceptual similarity

[1] Plaisted and O'Riordan generally express their theory in terms of "enhanced discrimination" rather than "reduced perceptual similarity". Although these two expressions are not identical, there is general agreement in the categorization literature that the perceptual *similarity* of two objects is monotonically related to the extent to which they can be *discriminated* (see, for example, Medin & Schaffer, 1978; Nosofsky, 1986, 1987; Shepherd, 1987). Put simply, an individual who finds it relatively easy to discriminate between two objects would also consider them relatively dissimilar. However, the concepts of discrimination and (dis)similarity are not used interchangeably in the categorization literature, and the theory described by Plaisted et al. (1998a) corresponds better to the categorization expression "reduced perceptual similarity" than to the term "enhanced discrimination". Because the techniques and theory employed in this article are based on standard paradigms used in perceptual categorization, we adopt the term "reduced perceptual similarity" for the remainder of the article.

hypothesis in terms of exemplar models of categorization (e.g., Medin & Schaffer, 1978; Nosofsky, 1986, 1990, 1992; Nosofsky & Palmeri, 1996). According to such models, people represent categories by storing individual exemplars with a label indicating their category membership. A novel item is categorized by computing its similarity to all the exemplars in memory, summing the similarities for each category and then assigning it to the category with the highest summed similarity to all items in memory. In fact, old items are also categorized in the same way, by computing similarities to all exemplars (including themselves). This means that items that are similar to members of another category are likely to be misremembered. Exemplar models therefore provide an explicit link between perceptual similarity, categorization, and generalization, and they allow predictions to be made about the effect of reduced perceptual similarity on categorization performance (see Hayes & Taplin, 1992, 1993a, 1993b, for a discussion of categorization from a developmental and clinical perspective).

One prediction that can be derived from the reduced similarity hypothesis is that people with autism will show a reduced prototype effect (Plaisted, 2001). The prototype of a category is the category member whose features represent the average of the group. As such, its summed similarity to all the category members is higher than that for any other item, and it is therefore more likely to be correctly classified than any other category member. However, if all the members of a category are seen as highly distinctive then this effect will be reduced because other group members will exert less of a reinforcing effect on the prototype.

Consistent with this prediction, Klinger and Dawson (2001; see also Klinger & Dawson, 1995) reported that children with autism showed a reduced prototype effect. Participants were taught to discriminate between two sets of cartoon animals. They were then presented with two novel members of one category, one of which was the prototype, and were asked to decide which was the best example of that category. Typically developing children reliably chose the prototype over the nonprototypical animal, but children with autism performed at chance levels. However, one potential problem with this study is that, although some dimensions differed within categories, others dimensions varied between categories only. For example, all members of one category had octagonal bodies, while members of other categories had bodies of different shapes. Participants with autism may therefore have learnt to distinguish between the categories by attending to such invariant features and would then have been unable to choose the most representative example of the category. This would have led to a reduced prototype effect but would not imply reduced perceptual similarity.

The current study also used a categorization paradigm to investigate the reduced perceptual similarity hypothesis. However, we avoided the difficulty suggested above by employing a design whereby participants were forced to attend to multiple dimensions of the stimuli in order to learn the category structure. In the training phase of our experiment, participants learned to classify 10 rectangles varying in height and width into two experimenter-determined categories, referred to as Categories A and B. Rectangles were chosen because they could vary along continuous dimensions, thus discouraging the use of verbalized rules. The stimuli were presented one at a time on a computer screen, and participants were given feedback informing them whether their classification was correct or not. Once participants had learned to correctly classify all 10 rectangles accurately, they proceeded onto a test phase in which they classified the same 10 rectangles together with 6 novel rectangles, but this time without any feedback.

Figure 1 shows the coordinates of the stimuli in terms of height and width, together with their appropriate classification in the training phase (either A or B). It can be seen that the majority of the Category B rectangles are located in the bottom left triangle of the rectangle space, while the A rectangles are in the top right triangle. However, there is one B rectangle (B6) that is the exception to the general rule because it is

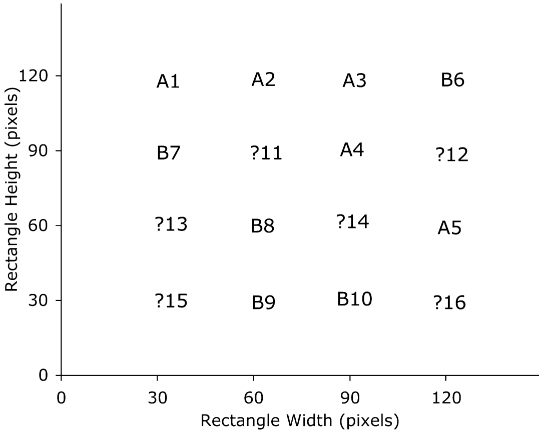

Figure 1. *Stimuli structure for the categorization task. Each label refers to a rectangle of a given height and width. Rectangles marked with an A or a B refer to items that were presented in the training phase (and belonged to category A or B respectively). Rectangles marked with question marks indicate rectangles seen only in the testing phase of the experiment. The numbers adjacent to the rectangles relate to the stimuli numbers shown in Figures 2 and 3. The rectangle B6 is surrounded by A category members and is therefore referred to as the exception item.*

located in the top right corner of the space and is consequently surrounded by Category A rectangles. We refer to this rectangle as the *exception item*, an important part of the design that we discuss below. Note that it is not possible to learn the correct labels of the rectangles by attending to one dimension of the stimuli only, whether this be height, width, area, or shape.[2]

Once participants had learned to correctly classify all 10 rectangles, they proceeded onto a test phase in which they classified the same 10 rectangles together with 6 novel rectangles, but this time without any feedback. According to exemplar models, classification of the exception item involves computing the similarity of the exception item to itself, to the Category A rectangles and to the remaining Category B rectangles. Because the similarity to Category A members is relatively high, participants are likely to make more errors in classifying the exception item. However, the extent of this effect will be determined by how much an individual is generally able to discriminate the different rectangles: If they have low discrimination abilities, then the Category A rectangles will be considered highly similar to the exception item, so the probability of misclassification will be relatively high. But if an individual can easily discriminate the rectangles, the interference from the surrounding A exemplars will be low, and performance on the exception item will be good. Thus, if people with autism judge the rectangles to be less similar than controls, then they should be better at discriminating the exception item in the testing phase of the experiment.

The experimental design also allowed quantitative modelling of categorization performance using the general context model (GCM; Nosofsky, 1986). The GCM is an exemplar-based model of categorization, which assumes that items are classified on the basis of their similarity to all items whose category membership is known. Similarity is defined using Shepard's (1987) universal law of generalization, which states that the similarity between two objects is determined by two factors: first the distance between the objects in psychological space, and second a scaling parameter representing a participant's level of memory sensitivity in discriminating among distinct exemplars. A high value means that the participant distinguishes easily among items in memory, while a low value means that exemplars are difficult to discriminate. The reduced perceptual similarity hypothesis therefore predicts higher values of the scaling parameter for individuals with autism than for controls.

We have so far emphasized predictions of the reduced perceptual similarity hypothesis for the testing phase of the experiment and not the training phase. This is because we wish all participants to be equated on the extent to which they know the training exemplars before being tested on

[2] If a participant tried to classify the rectangles on the basis of height alone, then it would not be possible to correctly classify rectangles B7 and A4 (amongst others) because they have the same height but are in different categories. Similar arguments hold for width (e.g., rectangles B9, B8, and A2), area (rectangles A1 and B8), and shape (B8 and A4 are both square).

generalization. By training participants until they reach a given criterion of accuracy, we eliminate the possibility that the different groups are generalizing from different knowledge bases. Nonetheless, it is possible to consider what predictions could be made from the reduced perceptual similarity hypothesis concerning category learning, as opposed to generalization.

Recall that participants who have high discrimination abilities would treat the rectangles as relatively distinct, individual items associated with the same category label. This implies that the exception item would be relatively easy to learn because there would be less interference that could arise from the surrounding rectangles of the opposing category. However, the normal rectangles would benefit from a high similarity because, overall, similar rectangles are placed within the same category. Thus, high discrimination would facilitate learning on the exception item but would harm learning on the normal rectangles. The reduced perceptual similarity hypothesis would therefore predict that the participants with higher functioning autism (HFA) would have less difficulty in acquiring the exception item than would controls, but more difficulty in learning the normal items. Predictions for the overall learning times are difficult to generate because there are both costs and benefits to high similarity, and the precise ratio cannot be determined in advance.

It is important to realize that there are many other factors aside from exemplar similarity that determine the speed with which participants acquire exemplars and the overall category structure. These include individual factors such as general confidence level and learning strategies, factors that are difficult to control for and likely to vary across our two groups. In general, learning a category is a far more complex cognitive activity than naming the exemplars, suggesting that the clearer test of reduced perceptual similarity would be the responses in the testing phase and not those from the training phase.

Fitting the GCM required that participants perform a large number of trials to eliminate as much noise as possible from their data. This consideration, together with the inherent difficulty of the task, entailed that it was not possible to test children or low-functioning individuals with autism. The participants in this study were therefore high-functioning adults with autism and nonautistic controls matched on verbal mental age and performance mental age.

Another requirement for modelling the data was that we needed to know how participants perceive the stimuli used in the experiment. Participants may view the stimuli in the same way as we have constructed them (as in Figure 1), or it may be that they perceive the rectangles in different way, by treating the rectangles as if they vary along only one dimension: area, for example. The GCM requires a set of psychological coordinates for the rectangles to describe categorization performance accurately. The most common approach to determining this set of coordinates is by performing a similarity ratings experiment using the stimuli that will be used in the principal experiment, in our case the 16 rectangles shown in Figure 1. This involves presenting participants with all possible pairs of stimuli and asking them to rate how similar they think each of the pairs are. These data are then analysed using a multidimensional scaling (MDS) analysis that converts the "distances" (similarity judgements) between stimuli into a "map" where each of the stimuli is given its own coordinates. The set of coordinates for all the rectangles is chosen so that the distances between stimuli on the map are as close as possible to the original distances given by participants. The resulting map is not restricted to two dimensions, but can be of any dimensionality from 1 to $N - 1$, where N is the number of different stimuli used in the experiment. Choosing the most appropriate number of dimensions for the space is generally done a priori or on grounds of parsimony, in a similar way to the approach taken in factor analysis. The final map provides a representation of the psychological space and can be used by the GCM. In order to determine the psychological coordinates of our stimuli, therefore, we performed a similarity ratings experiment with our HFA and our control participants, before commencing the categorization study (see Kruskal & Wish, 1978, and Young & Harris, 1994, for useful introductions to MDS).

In addition to being a useful step in the modelling of the GCM, this task also allowed investigation of potential differences between groups in the representation of the stimuli. In particular, participants might choose to encode the stimuli using spaces of differing dimensionality. For example, those in the HFA group may base their similarity judgements on only one dimension of the stimuli, whereas those in the control group might use two dimensions. Note, however, that it is not our aim to test the reduced perceptual similarity hypothesis using the similarity ratings task: Although we might expect the theory to predict lower similarity ratings overall for the HFA participants, differences that occur between the groups might also arise because of response biases that do not necessarily reflect underlying perceptual similarity. Because the effects of reduced perceptual similarity and response bias would be confounded, we restrict our investigation to differences involving the number of dimensions on which participants represent the stimuli. Response biases variation across individuals would not pose a problem for our analysis because individual ratings are normalized as part of the MDS algorithm.

EXPERIMENT

Method

Participants

A total of 12 high-functioning adults with autism were recruited via personal connections and local support groups and services. They had all been diagnosed by psychiatrists or clinical psychologists as having Asperger syndrome. However, it was not clear from the available information that early language development had been entirely normal in all cases as required by American Psychiatric Association (1994) and World Health Organization (1992) definitions of Asperger syndrome. The more open term higher functioning autism (HFA) was therefore used to describe these participants. All were given travel expenses and a gift voucher worth £10. A total of 17 controls were also recruited. These participants were undergraduate students at the University of Warwick and were paid £10 for taking part. Verbal mental age was assessed using the vocabulary and comprehension subtests of the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1986). Performance mental age was assessed using the picture completion and object assembly subtests. Participant details are summarized in Table 1. There were no significant group differences in performance mental age or verbal mental age (ts < 1), although the HFA group were significantly older than controls, $t(27) = 2.11$, $p = .043$.

Stimuli

Stimuli were 16 rectangles presented on the screen of a computer monitor. The outline of each rectangle was drawn in green single pixel lines on a black background. The dimensions of the rectangles represented the factorial combination of four heights and widths, which were 30, 60, 90, and 120 pixels. The same stimuli were used for the similarity ratings task and for the categorization experiment.

Table 1. *Participant details*

				Age[a]		*Verbal mental age*[a]		*Performance mental age*[a]	
	n	*No. male*	*No. female*	*M*	*Range*	*M*	*Range*	*M*	*Range*
HFA	12	10	2	30	20–62	27.10	17–36	19.5	7–31
Controls	17	6	11	21	19–45	27.12	20–35	21.6	17–31

[a]In years.

Procedure

Participants were tested individually in a soundproof booth with the experimenter in an adjoining room. The similarity ratings task was performed first, followed by the categorization task and finally the WAIS subtests.

In the similarity ratings task, participants were told that they would see pairs of rectangles presented on a computer screen and that they would have to judge how similar they thought each pair was. They were told to use the numbers 1–9 on the computer keyboard to make their judgement, 1 being the least similar, and 9 being the most similar. They then completed 480 trials in which they saw each of the possible rectangle pairs twice. A rectangle was never paired with itself, and the second time a pair was presented their positions to the left and right of the screen were changed. Rectangle pairs were presented in a different random order for each participant.

In the categorization task, participants were told that one of the experimenters had chosen some of the rectangles to be his, while the other experimenter had chosen some of the remaining rectangles. Participants were then told that they had to learn which rectangles belonged to each experimenter and that no one rectangle could belong to both experimenters. They were informed that they would receive feedback in the first part of the experiment (the training phase), but later on that feedback would disappear in the second phase (the testing phase). They were also told that they would see "new" rectangles in the testing phase and that they should classify these rectangles on the basis of the classifications that they had made in the training phase.

In each block of the training phase, the 10 training items were each presented once in a random order. A typical trial consisted of the presentation of a fixation point (1 second) followed by the rectangle appearing on the screen and remaining until the participant had made their response. All responses were made using a standard button-box, with two buttons corresponding to the two different categories of rectangle. Feedback was provided with a high beep if the participant responded correctly and a low beep if they responded incorrectly. Training continued until participants could correctly identify each of the 10 rectangles without error for four complete blocks, whereupon they moved onto the testing phase after a short break. There were 40 blocks in the test phase, with all 16 rectangles being presented once in a random order in each block, making a total of 640 trials. The procedure was the same as that in the training phase, apart from the absence of any feedback.

Results

Similarity ratings

The goal of this experiment was to provide a suitable set of coordinates for the GCM and to establish whether there might be any differences between the perceptual representations of individuals in the HFA and control groups. To this end, a type of MDS known as individual scaling (INDSCAL, Carroll & Chang, 1970) was applied, assuming ordinal data. The INDSCAL analysis results in a single solution for all participants together with a set of weights for each participant dictating the extent that the participant relies on each dimension. The GCM can then be applied using the average solution while differences between groups can be analysed using the dimensional weights.

We report the results of applying MDS assuming a one-, two-, or three-dimensional solution. The stress and R^2 figures that we report are the mean of each individual participant's fit to the single group solution. In the two-dimensional case, the INDSCAL analysis resulted in stress and R^2 values of .3 and .46 respectively, and for the three dimensional case they were .23 and .52. A replicated MDS analysis was applied to obtain a one-dimensional solution, with accompanying stress and R^2 values of .45 and .40, respectively (INDSCAL cannot be applied because there is only a single dimension). We did not try to select the solution with the most appropriate dimensionality because we were able to fit the GCM using each of the solutions and compare the results (see the Model Fitting section below).

We now examine whether there are differences between the two groups in how they perceive the rectangle space. One possibility is that participants in the two groups might be making their similarity judgements using different numbers of dimensions. For example, the HFA group might prefer to use only one dimension on which to base their similarity ratings, while the controls might prefer two dimensions. This hypothesis can be investigated by looking at the participant weightings on the dimensions, as a function of group. In addressing this question, we are assuming that at least some of our participants encoded the rectangles using two or three dimensions whereas, in fact, the R^2 results reported above do not lead us to reject the hypothesis that all participants used only a one-dimensional solution. We proceed with the analysis of higher dimensional solutions because individual variation in the dimensionality of encoding (the very question we are investigating here) contributes to overall variance and may have masked the extent to which it is possible to determine the best group solution.

To examine the extent to which different participants used a different number of dimensions, we analysed a transformation of the dimensional weights referred to as the *w*-score[3] (MacCallum, 1976; Young & Harris, 1994). The *w*-score indicates how far the individual's weightings differed from the group average: The more a participant relies on one dimension, the more extreme the weight ratio becomes and the higher their *w*-score. Consequently, if the HFA group were making their similarity judgements based on fewer dimensions than were the controls, their *w*-scores should be higher. For the two-dimensional solution, the mean *w*-scores were 0.242 (SD = 0.158) for the HFA group, and 0.156 (SD = 0.100) for the control group. A two-tailed, equal-variance t test revealed this difference to be narrowly nonsignificant, $t(27) = 2.00$, $p = .055$. Analysis of the weights for the three-dimensional solution revealed a similar story, with means of 0.27 for the HFA group and 0.2 for the controls, $t(27) = 1.6$, $p = .1$. These results provide some tentative support for the idea that the HFA group were making their similarity judgements on the basis of fewer dimensions than were the controls.[4] However, due to the difficulty of establishing the most appropriate group space, and the ambiguous results of the t tests, we refrain from drawing firm conclusions regarding differences in the representation of the rectangle space between groups.

Categorization task

For both training and testing phases, an arcsine transformation was carried out on all choice proportions to improve the conformity of the data to the standard assumptions of analysis of variance (ANOVA; Howell, 1997). In the training phase, 4 of the HFA participants failed to reach the criterion of four consecutive blocks of correct responses. Of these, 2 failed to perform above chance on a single block and were excluded from all further analyses.[5] The other 2 participants

[3] Referred to as the "weirdness" index by Young and Harris (1994).

[4] An alternative to analysing the participant weights might have been to fit two completely different solutions, one for the HFA group and another for the control group, and to examine the degree to which each solution fitted the data. One might have expected that, if the HFA group were basing their judgements on fewer dimensions, then the most optimal dimensionality for the HFA group would be lower than that of the controls. However, there are two problems with this method, which would make any result difficult to interpret. The first is that formal methods for deciding dimensionality, such as Lee's (2001) BIC, favour a low dimensional solution for data that has a high variance. Thus, different ideal dimensionalities could be the result of differences in the variance across groups, rather than a difference in the perceived rectangle space. The second problem is that even if the HFA solution requires a high-dimension solution, it could be because different participants within the group choose to base their judgements on different dimensions. Hence, a high-dimensional solution is required to account for between-participant variation in the choice of dimensions, even though any single participant might be best modelled with a low-dimensional solution. The INDSCAL technique that we used avoids both of these potential problems.

[5] The removal of these two HFA participants did not substantially alter the matching between our two groups. The new means and ranges for the HFA group were: 31.2 (20–62), 27.3 (20–35), and 22 (17–31) years for age, verbal mental age, and performance mental age, respectively. As before, there were no reliable differences between groups on the verbal and performance mental age measures, $t(25)$s < 1, ps $> .83$, and the difference on age remained, $t(25) = 2.6$, $p = .014$.

were very close to criterion (for example, getting one response wrong out of a run of 4 blocks) and proceeded onto the test phase after 100 training blocks. All analyses were therefore carried out both with and without these two noncriterion learners (NCLs). However, only the results including the NCLs are reported, unless there were qualitative differences between the conclusions of the two analyses. One HFA participant (not an NCL) dropped out three quarters of the way into the test phase because he was tired. His results were therefore based on the responses that were collected (27 blocks out of 40).

Training phase

Participants in the HFA group took more blocks than did controls to reach criterion in the training phase $t(23) = 2.74$, $p = .012$, with averages of 46 and 30 blocks, respectively. NCLs were not included in this comparison because the precise number of blocks required for them to reach criterion was unknown (although assuming learning times of greater than 100 blocks would increase the difference between the two groups). Turning to the proportion correct as the dependent measure, Figure 2 shows the accuracy of responses given to each rectangle in the training set (NCLs included). These results were subjected to an ANOVA with rectangle as a repeated measure and group as a between-subjects factor. The HFA group performed significantly worse than controls, $F(1, 27) = 4.43$, $p = .045$, reflecting

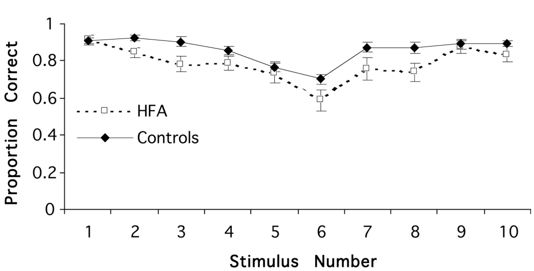

Figure 2. *Proportion correct during the training phase. Error bars are the standard errors for each stimulus, within developmental group. Stimulus numbers refer to the category structure shown in Figure 1, where 6 is the exception item.*

the greater number of blocks required to reach criterion. However, this effect became nonsignificant when the NCLs were removed, $F(1, 23) = 3.7$, $p = .067$. There was a significant effect of rectangle, $F(1, 9) = 12.07$, $p = .001$, demonstrating that the exception item was learnt with more difficulty than other rectangles. In fact, 7 out of the 10 participants in the HFA group and 11 of the 17 controls found the exception item (Stimulus Number 6) the most difficult to learn. There was no significant interaction between group and rectangle, $F(9, 243) = 1.15$. In summary, the HFA group found the category structure significantly more difficult to learn than did the controls, but there was no evidence of differences in the rate at which they learnt the exception item.

The above analysis was conducted on the data from the entire training phase, during which different participants received different amounts of training. However, it could be argued that the extra blocks of exemplars seen by the HFA group might obscure differences in discrimination or general learning strategy. For example, the HFA participants might have learned the exception item better than did controls earlier on learning but this effect could then have been obscured by requiring a large number of trials to learn the other items. To investigate this issue, we analysed the data from the first 25 blocks only. Up until this point, all participants were still in the training phase and had therefore received the same amount of training. We divided the 25 blocks into five divisions and, for each participant, found the mean accuracy for the exception item and for the normal rectangles. We averaged together responses to the normal rectangles, unlike in the analysis above, because we had far fewer data points than previously and we also had an extra factor, that of block. We then performed an ANOVA on these figures, with rectangle (normal and exception) and block (1 to 5) as repeated measures factors, and group (HFA and control) as a between-participants factor. The results of this analysis were very similar to results of the analysis carried out using the data from the entire training set. We found that both groups learned the exemplars better over time,

$F(4,\ 100) = 32.08$, $MSE = 0.29$, $p < .0005$, and that classification of the exception rectangle was poorer than that for the others, $F(1, 25) = 29.40$, $MSE = 0.42$, $p < .0005$. This effect was present in both the control participants, $F(1, 16) = 19.18$, $MSE = 0.40$, $p < .0005$, and the HFA group, $F(1, 9) = 11.67$, $MSE = 0.45$, $p = .008$. The difference between the exception item and the normal rectangles diminished over blocks, $F(4,\ 100) = 5.14$, $MSE = 0.30$, $p = .001$. The HFA group performed worse than controls overall, $F(1, 25) = 3.6$, $MSE = 1.24$, $p = .07$, and this difference diminished over learning, $F(4,\ 100) = 2.34$, $MSE = 0.29$, $p = .061$. However, there was no interaction of group with rectangle, $F(1,\ 25) = 0.045$, $MSE = 0.42$, $p = .834$, nor was there an interaction of group, rectangle and block, $F(4,\ 100) = 0.771$, $MSE = 0.303$, $p = .547$. As we found when we analysed the complete training set, the HFA group appeared to learn more slowly overall, but there was no evidence of any learning differences related to the exception item.

Test phase

Figure 3 shows the proportion of correct responses in the test phase to the 10 rectangles that participants had learnt to classify during the training phase. Performance on the exception item (Stimulus 6) was clearly worse than that on the other nine items, despite the fact that all items had been learnt to criterion in the training phase.

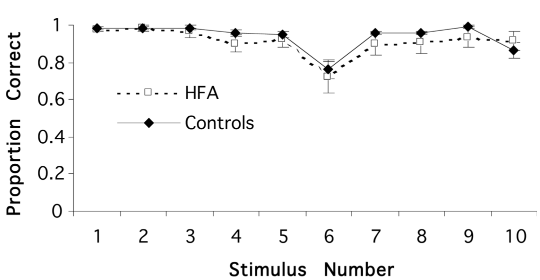

Figure 3. *Proportion correct for responses to training items, during the testing phase. Stimulus numbers refer to the category structure shown in Figure 1, where 6 is the exception item. Error bars are the standard errors for each stimulus, within developmental group.*

This observation was confirmed by performing a repeated measures ANOVA on the proportion correct with group and rectangle as factors. A significant main effect of rectangle was found, $F(9, 225) = 17.12$, $MSE = 0.1$, $p < .001$, indicating that the exception item was reliably different to other items (see Figure 3), but there was no main effect of group, $F(1, 25) = 0.9$, $MSE = 0.56$, $p = .352$. Importantly, there was no significant interaction between group and rectangle, $F(9, 225) = 1.69$, $MSE = 0.1$, $p = .094$, a finding that fails to support the reduced perceptual similarity hypothesis.

Despite our failure to find a reliable interaction, we analysed performance between groups on the exception item in particular because the reduced perceptual similarity hypothesis makes a priori predictions regarding this item. Mean proportion correct for the HFA group was .73 ($SD = 0.10$), while for the controls it was .78 ($SD = 0.050$). Thus, the means were in the opposite direction to that predicted by the reduced perceptual similarity hypothesis. Analysis of confidence intervals demonstrated that we can be 95% certain that the mean HFA correct score cannot be more than .014 greater than that of the control group, assuming $t(25)$, one-tailed, because the reduced perceptual similarity hypothesis predicts $M_{\mathrm{HFA}} \gg M_{\mathrm{control}}$. Thus, these results suggest that any effects of reduced perceptual similarity are negligible in such a task as ours.

We also conducted a correlational analysis involving the factors of age, verbal mental age, and performance mental age. This analysis revealed that classification of the exception item was not associated with age ($r = .194$, $p = .33$) nor with verbal mental age ($r = .196$, $p = .53$), but was positively correlated with performance mental age: $r = .51$; $Z(27) = 2.748$; $p = .006$. A one-way analysis of covariance (ANCOVA) was therefore performed with classification of the exception item as the dependent variable, group as the between-subjects factor, and performance mental age as the covariate. The main effect of performance mental age approached significance, $F(1,\ 23) = 4.00$, $p = .058$, and there was a significant interaction between the group and performance mental age, $F(1, 23) = 4.6$, $p = .043$. Crucially, however, once the variance due to performance mental age had been factored out, control participants showed reliably better responses on the exception item, $F(1, 23) = 4.7$, $p = .041$. A reliable interaction between the two groups of participants and performance mental age was present, $F(1, 23) = 4.6$, $p = .043$, although the main effect of performance mental age was narrowly nonsignificant, $F(1, 23) = 4.00$, $p = .058$.

In summary, the results indicate that participants in the HFA group did not perform more accurately than those in the control group. If anything, the small effects present in the ANCOVA analysis point to an advantage for the control group—effects that are in the opposite direction to those predicted by the reduced perceptual similarity hypothesis.

Figure 4 shows the classification of the six novel rectangles in the test phase as indexed by the proportion of A responses. Results were subjected to an ANOVA with rectangle as a repeated measure and group as a between-subjects factor. There were no reliable effects involving group (Fs < 1).

Model fitting

The results from the test phase were modelled using the GCM (Nosofsky, 1986). This model assumes that exemplars are represented as points in multidimensional psychological space, with exemplars from a particular category tending to cluster together because they generally have similar values on each dimension. The coordinates

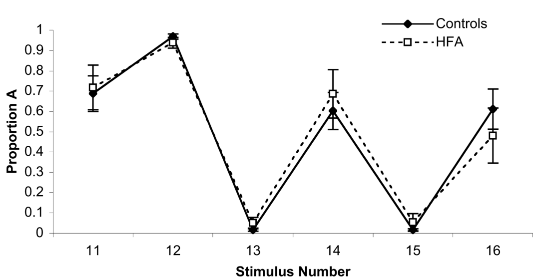

Figure 4. *Proportion of 'A' responses for the rectangles presented in the test phase only. Error bars are the standard errors for each stimulus, within developmental group.*

of the exemplars are the psychological coordinates derived from the MDS analysis of similarity ratings.

The distance, d_{ij}, between two items X_i and X_j is defined by the equation

$$d_{ij} = \left[\sum_m w_m |x_{im} - x_{jm}|^r \right]^{1/r} \qquad (1)$$

where x_{im} and x_{jm} are the coordinates of the two items on dimension m. The r parameter dictates the metric of the space. If r is equal to 1 then the distance is determined by summing the differences between the coordinates of the two objects on each dimension, whereas if r is equal to 2 (Euclidean space) then distance is simply the shortest straight line between the two points. The similarity, s_{ij}, between X_i and X_j is then defined as an exponential function of the distance between them:

$$s_{i,j} = \exp(-c \cdot D_{i,j}) \qquad (2)$$

where c is a scaling parameter that determines how easy the individual objects are to discriminate. Figure 5 shows how similarity varies as a function

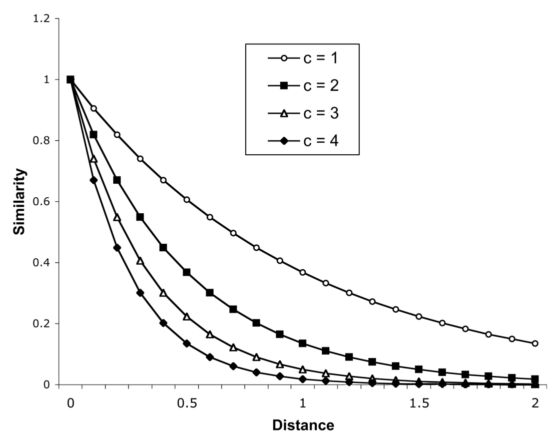

Figure 5. *Similarity as a function of distance and the value of the* c *parameter. At high values of c, objects are less similar than at low values.*

of c and the distance between two items as determined by the coordinates of the object in psychological space. When c is high, the function drops off sharply so that at a given distance apart, two items are relatively dissimilar and can be easily discriminated. Conversely, when c is low, the same two items appear more similar and are relatively difficult to discriminate. The c parameter is generally a free parameter estimated by minimizing the difference between the model's predictions and an individual participant's responses. Put in these terms, the c parameter can be seen as a useful way of measuring a participant's sense of perceptual similarity. As such, the reduced perceptual similarity hypothesis would predict increased c parameters in the HFA group relative to controls.

The final stage of the model involves assigning the item to a particular category. In the GCM, the probability that the object, X_i, is assigned to a particular category C_k is given by the sum of the similarities between X_i and each item in C_k divided by the summed similarity of X_i, to all items in memory.

$$p(C_k \mid X_i) = \frac{\sum_{j \in k} s_{ij}}{\sum_{j=1}^{N} s_{ij}} \qquad (3)$$

In summary, the algorithm first calculates the distance between the unknown item and the known exemplars using Equation 1, then transforms these distances into similarities using Equation 2, and finally calculates the probability that the object belongs in a particular category using Equation 3.

The GCM was fitted to the data by adjusting the c parameter value to minimize the summed square error between each participant's responses and the model predictions, resulting in a single c parameter value for each participant. Because there were no a priori assumptions regarding the most appropriate MDS solution, or the most appropriate GCM distance metric, this analysis was performed for all possible combinations of these factors.

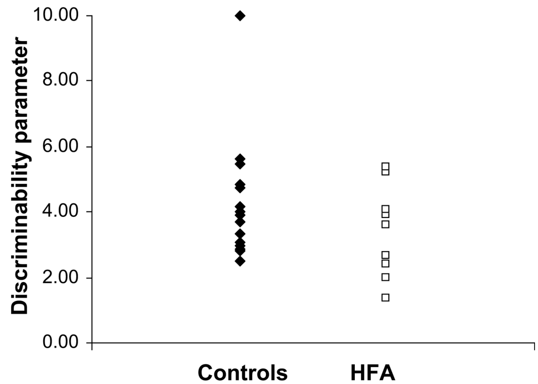

Figure 6. *The c parameter scores after fitting the GCM to the categorization response, using the objective MDS coordinates.*

We first used the MDS solutions as the set of exemplar coordinates for the GCM. All three solutions were tested, using both $r = 1$ and $r = 2$, which resulted in a value of R^2 that varied between .56 and .70. However, there was no evidence for any group differences in c parameter values. In fact, the GCM demonstrated the best fits to the data using the objective coordinates of the rectangles (i.e., their heights and widths) rather than psychological coordinates derived from MDS. With the GCM distance metric, r, set to 1, the R^2 value was .74, and when r was set to 2, the R^2 value was .75. Figure 6 shows the c parameters for each participant with $r = 2$ (results were similar with $r = 1$).[6] The reduced perceptual similarity hypothesis predicted that the HFA group would have higher c parameters than those in the control group. In fact, the difference was in the other direction, although not reliably so, $U'(17) = 161$, $p = .29$. Furthermore, all the participants in the HFA group appear to have c parameters well within the range of the control participants, indicating that individual differences are not obscuring a group effect.

[6] One participant in the control group had an extremely high c value (the maximum possible value tested by our optimization algorithm). This was because he had a near-perfect score for the training items during the testing phase, which the GCM fits best by using an infinitely high c parameter value.

These results do not provide evidence in support of the reduced perceptual similarity hypothesis.

One aspect of our results that might seem unusual is that we obtained better model fits with the objective solution than with the MDS solutions. We believe this stems from two factors of our design. First, because we were using a clinical sample, there was likely to be considerable individual variation in the similarity judgements. This would have resulted in a noisy set of coordinates from the INDSCAL analysis. If the "true" perceptual representation resembled the objective coordinate set more than the results of the INDSCAL analysis, then a better fit would be expected from the GCM with the objective set. Second, the GCM was modelled on responses that took place after the training phase of the experiment. If there were participants who chose to rely on only one dimension during the similarity ratings task, they would have had to change their representational coordinates during training, in order to learn the category structure. Thus, the true perceptual space may well have changed from that described by the MDS analysis to the objective space by the time participants made their responses during the testing phase.

Discussion

The current study investigated the hypothesis that autism is associated with a reduced sense of perceptual similarity. High-functioning individuals with autism and nonautistic controls first performed a similarity ratings task, designed to determine the psychological coordinates of the rectangle stimuli that were used in the subsequent categorization task. Analysis by MDS revealed a trend suggesting that individuals in the HFA group represented the stimuli on fewer psychological dimensions than did controls, although this pattern was not significant. Participants then performed a categorization task in which they learnt to classify rectangles into two arbitrarily defined categories. Individuals with autism took reliably longer to learn the category structure during the training phase of this task. However, the main focus of this study was performance on the test phase of this task in which participants were required to categorize the learned stimuli without feedback and generalize their responses to stimuli that had not previously been categorized. As expected, performance was relatively poor on an exception item that was similar to members of the opposite category but, contrary to predictions, there was no evidence that this effect was reduced among members of the HFA group. Moreover, generalization of responses to novel stimuli was similar in both groups.

Similarity ratings task

The main objective of the similarity ratings task was to provide psychological coordinates for modelling of the categorization task. Nevertheless, the results of the MDS analysis suggested potentially interesting differences between individuals with and without autism in terms of their initial representations of the stimuli. Specifically, participants in the HFA group showed a trend towards higher w-scores in the INDSCAL analysis, indicating that they might be basing their ratings of the stimuli on fewer dimensions than were controls. However, given that the effect was nonsignificant, this result should of course be treated with caution. Furthermore, the current data do not demonstrate that the control participants encoded the data using more than one dimension; a finding that would be a necessary precursor for using the w-score analysis as evidence that participants with autism were representing the data on fewer dimensions than were controls. Moreover, the current study was not specifically designed to test the hypothesis that individuals represent stimuli on fewer dimensions than were controls, and, consequently, all interpretation of the MDS analysis is post hoc. Future experiments could investigate this issue using the same methodology, although we would recommend collecting more data points per participant to reduce within-participant noise and using high-dimensional stimuli, rather than rectangles, to test the integration ability of people with autism more fully.

Category learning

Although category learning was not the main focus of this study, we note that the training phase of the categorization task provided the first reported investigation of the process of category-learning abilities of people with autism under controlled conditions. Perhaps unsurprisingly, given that individuals with autism appear to have difficulties forming categories, the HFA group took significantly longer than the control group to learn the category structure. Advocates of the reduced perceptual similarity hypothesis may be tempted to explain this in terms of difficulties in noticing the similarities between category exemplars. This would imply that for the HFA group, the costs of not being able to group together similar rectangles outweighed the benefits of learning the exception item relatively easily. However, both groups found that the exception item was reliably more difficult to learn than the other rectangles, and there were no differences in the rate in which different groups learned the different types of rectangle. Thus, we found no evidence to suggest that a reduced perceptual similarity hindered learning of the category structure.

The fact that the exception item produced most errors during training means that participants were required to change their initial responses to the exception item in order to learn the category structure. There is considerable evidence that people with autism have specific difficulties on tests of executive function when they are required to change the dimensions of a stimulus to which they are attending (e.g., Ciesielski & Harris, 1997; Hughes, Russell, & Robbins, 1994; Ozonoff & Jensen, 1999), and, arguably, such difficulties could explain the relatively slow learning of the category structure. However, it is important to note that, whereas tests of executive functioning require participants to inhibit responses that have previously been reinforced, in our task the correct response (and feedback) to the exception item was always the same. Thus, if executive deficits are to explain poor category learning then one must assume that individuals in the HFA group had difficulty inhibiting the exception item response because of the competition from the similar rectangles. This suggestion appears plausible but rests on the assumption that individuals in the HFA group are sensitive to the similarity of the surrounding items—an assumption that is at odds with the reduced perceptual similarity hypothesis.

Finally, it is possible that participants in the HFA group started the categorization task with inappropriate assumptions about the best way of representing the rectangles, and this hindered their learning because they had to change the representation during the learning process. This possibility is compatible with the idea, discussed above, that individuals in the HFA group initially represented the stimuli on fewer dimensions than did controls. Clearly, further research with paradigms specifically designed to investigate category learning will be necessary to distinguish between these different explanations for slow category learning in autism.

Categorization and generalization

As expected, performance in the testing phase was reliably poorer on the exception item than on the other rectangles, despite the fact that participants had learnt to correctly classify the exception item in the training phase of the experiment. This presumably was a result of interference from surrounding rectangles (see Figure 1). However, the reduced similarity hypothesis predicted that those in the HFA group would perform more accurately than controls on this exception item because they would be less influenced by the surrounding rectangles. In the event, there was no significant advantage for the HFA group.

The results of the test phase were modelled using the GCM (Nosofsky, 1986), which allowed us take into account responses to all the items, rather than just the exception item, and provided us with a measure of individual participant performance. The reduced similarity hypothesis predicted that individuals in the HFA group would have higher scaling parameters (i.e., c parameters) than would the controls. Again, however, the trend was in the opposite direction—there were no differences between the groups, and no individual participant in the HFA group differed from the sample of controls.

Overall, therefore, the results failed to support the reduced perceptual similarity hypothesis and contrast with a number of other recent findings. There are a number of potential explanations for this. First, it might be argued that there was insufficient statistical power in the design. However, participants were tested 40 times each on the exception item with little evidence of ceiling or floor effects that might have masked an advantage for the HFA group. Furthermore, confidence intervals for performance on the exception item indicated that any effect of reduced perceptual similarity is likely to be very small. A second potential criticism concerns the matching of controls: The mean age of the HFA group was higher than that of the control group, and performance mental age was lower than that of the control group. However, there was no evidence for a correlation between age and classification of the exception item, and controlling for performance mental age by covariation resulted in an advantage for the control group on the exception item. This latter result is in the reverse direction to the reduced perceptual similarity hypothesis.

A third and more theoretical concern might be that exemplar models are not appropriate to apply to this experimental design and that it is invalid to assume, for example, that performance on the exception item is determined by perceptual similarity, or that the results can be modelled with the GCM. The most plausible alternative model would be some form of rule-plus-exception model (e.g., Erickson & Kruschke, 1998; Nosofsky & Palmeri, 1998), which would take the form of a rule (e.g., "respond 'A' if the rectangle is in some corner of the space, 'B' if it is in another corner"), but with some rectangles remembered on an individual basis. However, these models still require an aspect of generalization around the exception item to account for the fact that it is remembered significantly less well than other old items during the testing phase (both in the control participants and in the HFA group). The task therefore assesses perceptual similarity even under the assumptions of a different model. Consequently, even if the modelling of the data using the GCM is not appropriate, the standard inferential statistics used to analyse responses to the exception item still provide evidence against reduced similarity.

A further issue concerns the generalizability of the current results. The complexity of the task and the large number of trials that had to be completed combined to ensure that it was only possible to test high-functioning adults with autism. One possibility is that reduced perceptual similarity is only found in younger or less able individuals with autism, and this could potentially account for the discrepancy between the current results and the reduced prototype effect demonstrated by the children with autism in the study conducted by Klinger and Dawson (1995, 2001). While this possibility cannot be firmly ruled out, our results do strongly suggest that findings from other studies with high-functioning adults with autism, such as the reduced transfer effect noted by Plaisted et al. (1998b) and the relatively good performance on the embedded figures task reported by Jolliffe and Baron-Cohen (1997), cannot be explained purely in terms of reduced perceptual similarity, as has been previously argued (Plaisted, 2001; Plaisted et al., 1998b).

Finally, we note that our findings are at odds with those of Bonnel et al. (2003), who recently found evidence for enhanced sensitivity to pitch in people with HFA—a result consistent with the reduced perceptual similarity hypothesis. One obvious difference between the studies is the type of stimuli used. Thus, it is possible that reduced perceptual similarity in autism is restricted to the auditory domain. A second potentially crucial difference is in the dimensionality of the stimuli; the current study employed multidimensional rectangles, whereas Bonnel et al. (2003) used stimuli that varied on a single dimension. It is possible that difficulties in integrating dimensions in the current study could mask otherwise superior discrimination abilities. However, it is important to note that by the time our participants had completed the training phase of our experiment, they had succeeded in integrating the multidimensional stimuli. Clearly, further research using simple and complex stimuli from different modalities is required to determine the strengths and weaknesses in categorization and discrimination abilities in autism.

The results of the current study provided no evidence to support the hypothesis that autism is associated with a reduced sense of perceptual similarity. Instead, we propose the novel hypothesis that individuals with autism have a tendency to represent objects on fewer dimensions than do typically developing individuals, as we suggested above when discussing the similarity ratings task. This "reduced dimensions" hypothesis could explain the reduced prototype effect in autism reported by Klinger and Dawson (1995, 2001). It was suggested that individuals with autism in that experiment were not able to choose the most representative exemplar from the category because they had not coded the dimensions that defined the prototype. This situation may have arisen because children with autism chose to code the objects on the fewest dimensions possible to perform the training task—that is, on the between-category dimension—whereas controls chose to represent the objects more completely. The reduced dimensions hypothesis could also explain why the HFA group required more blocks to learn the category structure than did controls in this experiment—the HFA group would have had to change their perceptual space in order to learn to classify the rectangles, whereas the controls were already attending to the appropriate dimensions (see Palmeri & Nosofsky, 2001, for evidence that the perceptual space can be different after a categorization task). More speculatively, the hypothesis could also explain some of the difficulties in generalization in everyday situations that Plaisted (2001) has attributed to reduced perceptual similarity. Objects and situations in real life are typically complex and multidimensional. If one focuses on a very limited aspect of a stimulus (such as the colour of an object) then there is a high probability that this will be a superficial rather than a defining feature of the object. Subsequent encounters with other superficially different members of the same category will then be treated as an entirely novel experience.

GENERAL CONCLUSIONS

The experiments presented in this paper were designed to test the hypothesis that people with autism have reduced perceptual similarity. We found no evidence in support of this hypothesis. In addition to our principal finding, we have made several other contributions. First, we have introduced a precise definition of reduced perceptual similarity based on computational models from the study of categorization. Second, we have provided evidence that people with autism may have difficultly in learning categories. Third, we have suggested a novel hypothesis concerning the cause of some of the generalization difficulties in autism: that of coding objects on fewer dimensions than controls in circumstances in which such a strategy can be effective. The results of the current study are consistent with this hypothesis but it clearly demands further investigation.

Original manuscript received 13 May 2004
Accepted revision received 1 October 2004
PrEview proof published online 29 August 2005

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

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

Brian, J. A., & Bryson, S. E. (1996). Disembedding performance and recognition memory in autism/PDD. *Journal of Child Psychology and Psychiatry, 37*, 865–872.

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart–Young" decomposition. *Psychometrika, 35*, 283–319.

Ciesielski, K. T., & Harris, R. J. (1997). Factors related to performance failure on executive tasks in autism. *Child Neuropsychology, 3*, 1–12.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General, 127*, 107–140.

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

Frith, U., & Happé, F. (1994). Autism: Beyond "theory of mind". *Cognition, 50*, 115–132.

Happé, F. (1999a). Autism: Cognitive deficit or cognitive style? *Trends in Cognitive Science, 3*, 216–222.

Happé, F. (1999b). Understanding assets and deficits in autism: Why success is more interesting than failure. *Psychologist, 12*, 540–546.

Hayes, B. K., & Taplin, J. E. (1992). Developmental changes in categorization processes: Knowledge and similarity-based models of categorization. *Journal of Experimental Child Psychology, 54*, 188–212.

Hayes, B. K., & Taplin, J. E. (1993a). Developmental differences in the use of prototype and exemplar-specific information. *Journal of Experimental Child Psychology, 55*, 329–352.

Hayes, B. K., & Taplin, J. E. (1993b). The development of conceptual knowledge in children with mental retardation. *American Journal on Mental Retardation, 98*, 293–303.

Howell, D. C. (1997). *Statistical methods for psychology* (4th ed.). Belmont, CA: Wadswoth Publishing Company.

Hughes, C., Russell, J., & Robbins, T. R. (1994). Evidence for executive dysfunction in autism. *Neuropsychologia, 32*, 477–492.

Jolliffe, T., & Baron-Cohen, S. (1997). Are people with autism and Asperger syndrome faster than normal on the embedded figures test? *Journal of Child Psychology and Psychiatry, 38*, 527–534.

Klinger, L. G., & Dawson, G. (1995). A fresh look at categorization abilities in persons with autism. In E. Schopler & G. B. Mesibov (Eds.), *Learning and cognition in autism. Current issues in autism* (pp. 119–136). New York: Plenum Press.

Klinger, L. G., & Dawson, G. (2001). Prototype formation in autism. *Development and Psychopathology, 13*, 111–124.

Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling.* Beverly Hills, CA: Sage Publications.

Lee, M. D. (2001). Determining the dimensionality of multidimensional scaling representations for cognitive modeling. *Journal of Mathematical Psychology, 45*, 149–166.

MacCallum, R. C. (1976). Transformation of a three-mode multidimensional scaling solution to INDSCAL form. *Psychometrika, 41*, 385–400.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85*, 207–238.

Mottron, L., & Burack, J. A. (2001). Enhanced perceptual functioning in the development of autism. In J. A. Burack, T. Charman, N. Yirmiya, & P. R. Zelazo (Eds.), *The development of autism: Perspectives from theory and research* (pp. 131–148). Hillsdale, NJ: Erlbaum.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General, 115*, 39–57.

Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 87–108.

Nosofsky, R. M. (1990). Relations between exemplar-similarity and likelihood models of classification. *Journal of Mathematical Psychology, 34*, 393–418.

Nosofsky, R. M. (1992). Exemplars, prototypes, and similarity rules. In A. M. Healy, S. F. Kosslyn, & R. M. Shiffrin (Eds.), *From learning theory to connectionist theory: Essays in honor of William K. Estes* (pp. 149–167). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Nosofsky, R. M., & Palmeri, T. J. (1996). Learning to classify integral-dimension stimuli. *Psychonomic Bulletin & Review, 3*, 222–226.

Nosofsky, R. M., & Palmeri, T. J. (1998). A rule-plus-exception model for classifying objects in continuous-dimension spaces. *Psychonomic Bulletin & Review, 5*, 345–369.

O'Riordan, M., & Plaisted, K. (2001). Enhanced discrimination in autism. *Quarterly Journal of Experimental Psychology, 54A*, 961–979.

O'Riordan, M. A., Plaisted, K. C., Driver, J., & Baron-Cohen, S. (2001). Superior visual search in autism. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 719–730.

Ozonoff, S., & Jensen, J. (1999). Brief report: Specific executive function profiles in three neurodevelopmental disorders. *Journal of Autism and Developmental Disorders, 29*, 171–177.

Ozonoff, S., & Miller, J. N. (1995). Teaching theory of mind: A new approach to social skills training for individuals with autism. *Journal of Autism and Developmental Disorders, 25*, 415–433.

Palmeri, T. J., & Nosofsky, R. M. (2001). Central tendencies, extreme points, and prototype enhancement effects in ill-defined perceptual categorization.

Quarterly Journal of Experimental Psychology, 54A, 197–235.

Plaisted, K. (2001). Reduced generalization in autism: An alternative to weak central coherence. In J. A. Burack, T. Charman, N. Yirmiya, & P. R. Zelazo (Eds.), *The development of autism: Perspectives from theory and research* (pp. 149–169). Hillsdale, NJ: Erlbaum.

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

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

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

Shah, A., & Frith, U. (1993). Why do autistic individuals show superior performance on the block design task? *Journal of Child Psychology and Psychiatry, 34*, 1351–1364.

Shepard, R. (1987). Towards a universal law of generalization for psychological science. *Science, 237*, 1317–1323.

Swettenham, J. (1996). Can children be taught to understand false belief using computers? *Journal of Child Psychology and Psychiatry, 37*, 157–165.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327–352.

Tymchuk, A. J., Simmons, J. Q., & Neafsey, S. (1977). Intellectual characteristics of adolescent childhood psychotics with high verbal ability. *Journal of Mental Deficiency Research, 21*, 133–138.

Wechsler, D. (1986). *Wechsler adult intelligence scale—revised, UK edition.* Windsor, UK: NFER-Nelson.

Wing, L. (1996). *The autism spectrum: A guide for parents and professionals.* London: Constable.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 419–433.

World Health Organization. (1992). *International classification of diseases and disorders* (10th ed.). Geneva, Switzerland: Author.

Young, F. W., & Harris, D. F. (1994). Multidimensional scaling. In M. J. Norusis (Ed.), *SPSS professional statistics 6.1.* Chicago: SPSS, Inc.