J Autism Dev Disord (2008) 38:1721–1730
DOI 10.1007/s10803-008-0557-7

ORIGINAL PAPER

When Prototypes Are Not Best: Judgments Made by Children with Autism

Catherine J. Molesworth · Dermot M. Bowler · James A. Hampton

Published online: 10 April 2008

Abstract The current study used a factorial comparison experimental design to investigate conflicting findings on prototype effects shown by children with autism (Klinger and Dawson, Dev Psychopathol 13:111–124, 2001; Molesworth et al., J Child Psychol Psychiatry 46:661–672, 2005). The aim was to see whether children with high-functioning autism could demonstrate prototype effects via categorization responses and whether failure to do so was related to difficulty understanding ambiguous task demands. Two thirds of the autism group did show an effect. The remainder, a sub-group defined by performance on a control task, did not. The discussion focuses on the influence of heterogeneity within the autism group and the ability to resolve ambiguity on task performance. Finally, an alternative experimental design is recommended for further research into these issues.

Keywords Autism · Asperger syndrome · Categorization · Concepts · Heterogeneity · Prototype

Introduction

Autism and Asperger syndrome are developmental disorders that are diagnosed in the presence of impaired communication and social interaction as well as stereotypies, repetitive behaviors, or restricted interests. There is considerable overlap between the two conditions, however, clinically significant delays in language and in general cognitive development may be a feature of autism, but not of Asperger syndrome, DSM-IV (American Psychiatric Association 1994). In addition to these diagnostic features, individuals on the autism spectrum are characterized by atypicalities in learning and memory. Several authors (e.g. Klinger and Dawson 1995, 2001; Plaisted 2001; Tager-Flusberg 1985a) have suggested that these point to underlying difficulties with conceptual representation and categorization. For example, individuals with autism have trouble generalizing from old previously learnt material to novel information. Both children with high functioning autism (HFA) and children with low functioning autism (LFA) were less likely than comparison children to apply social training to real life social situations (Ozonoff and Miller 1995; Swettenham 1996).

The fact that individuals with autism tend not to use conceptual knowledge to aid memory represents further evidence of atypical categorization processes. LFA children, in contrast to comparison groups, are less likely to aid free recall by grouping exemplar information into categories (Hermelin and O'Connor 1970; Minshew et al. 1992; Tager-Flusberg 1991). This tendency has been identified also in adults with Asperger syndrome (Bowler et al. 1997, 2000).

Some researchers, however, have found no difficulties with categorization. For example, Ungerer and Sigman (1987) found no difference between LFA children and comparison children on the ability to categorize on a single basis (e.g. color or form). Tager-Flusberg (1985a, b) also found that LFA children showed comparable performance to comparison children in the ability to categorize exemplars into basic level categories (e.g. *boat*, *bird*) and superordinate categories (e.g. *food*, *tool*).

Two authors, Klinger and Dawson (1995, 2001) characterized this mixed pattern of categorization performance in a particular way: they suggested that individuals with

C. J. Molesworth (✉) · D. M. Bowler · J. A. Hampton
Department of Psychology, City University, Northampton Square, London EC1V 0HB, UK
e-mail: c.j.molesworth@googlemail.com

autism had trouble with a specific type of concept formation. A brief summary of the relevant theories of conceptual representation will follow before returning to Klinger and Dawson's account.

The classical view of concepts dominated the field of concepts until the 1970s. (See Hampton 1997 and Murphy 2002, for a review.) This view held that a concept is represented by shared properties that are individually necessary and jointly sufficient to define the concept. For example, the defining properties of a square are that the item has a closed figure, four sides, sides of equal length, and equal angles. Categorization was thought to proceed by means of simple *if… then* rules. However, by the mid 1970s, it had become apparent that such a view does not describe adequately how many real-world categorization decisions are made. For instance, people perceive typicality differences when making category membership decisions and tend to agree on which items are more typical than others. For example, apples are rated as a better example of fruit than water melons (Rosch 1975a). The classical view, under which category membership is regarded as either present or absent, does not predict such typicality effects.

An alternative to the classical view was the idea that many categories are represented by prototypes (best examples of the categories) and that these provide a summary of information in the category (Rosch 1975b). From this viewpoint the categorization of novel exemplars is carried out on the basis of how similar an exemplar is to the relevant category prototype; the greater this similarity the greater the probability of category membership (Rosch et al. 1976). Similarity also determines prototype effects. These can be observed in recognition memory where individuals tend to display false recognition to a previously unstudied prototype. Also characteristic of the effect is the fact that the degree of similarity between the exemplar and the prototype is reflected in recognition levels: the higher the similarity, the more likely a positive recognition response (Cabeza et al. 1999; Omohundro 1981; Solso and McCarthy 1981). A similar prototype effect has been observed using categorization. Unstudied prototypes are categorized with an accuracy that is at least equal to that of previously studied but less typical exemplars (Metcalfe and Fisher 1986; Posner and Keele 1968).

Klinger and Dawson (1995, 2001) drew upon clinical observation and empirical evidence to argue that individuals with autism behaved in the manner predicted by the classical model of concepts (using rule-defined concepts) because of difficulty abstracting and using prototypes (summary representations of categories). In support of this dissociation, Klinger and Dawson observed that children with autism are able to infer rules during the Wisconsin Card Sorting Test and similar set shifting tasks (Bennetto et al. 1996; Berger et al. 1993; Hughes et al. 1994). Additionally, individuals with autism persisted in rule-use on occasions where such a rigid approach is sub-optimal and where a prototype-based form of categorization would be more appropriate. For example, these authors reported the frustration of a father who tried to warn his adolescent autistic son not to interact with strangers. The problem was that his son kept on asking for a set of criteria that were necessary and sufficient for this concept.

Klinger and Dawson (2001) tested their account by comparing the two forms of categorization, rule-based versus prototype-based, for an LFA group and two comparison groups, one each of Down syndrome and of typical development. The experimental stimuli comprised categories of schematic animals. Each category was organized around a central prototype that possessed features (e.g. tails) that were a mean size of those possessed by other category members. Categorization performance was tested by asking each participant to select a member of a named category (e.g. *Mip)* from a target-lure pair immediately following a familiarization phase with that category. If participants responded that both target and lure were category members, they were instructed to select the best one. During rule-based conditions only, all category members possessed a feature that defined category membership such as a long foot, for example. Both target and lure were identical save for the presence or absence of the defining feature. During prototype-based conditions, the target was the category prototype, and had not been seen before. The lure was a novel composite: a category member that possessed individual features that had been seen in the study phase but in a novel combination. Only the typically developing group behaved as if they had abstracted prototype representations by selecting the target, the prototype, at levels significantly above chance. By contrast, in the rule-based conditions, all participant groups selected the target at levels that were above chance. The authors' main conclusion was that individuals with autism and Down syndrome had difficulties with prototype formation.

In contrast, we found that HFA children demonstrated full prototype effects via recognition responses (Molesworth et al. 2005). This discrepancy could be attributable to methodological differences between the two experiments. Our clinical group was more able than the ones participating in Klinger and Dawson's study. Any influence that ability exerts on performance is likely to be indirect. Existing evidence seems to suggest that prototype formation per se is a fundamental learning process unaffected by developmental variables or level of intelligence (Molesworth et al. 2005). Prototype effects have been observed in infants (Younger 1990) and individuals with organic developmental delay (Hayes and Conway 2000; Hayes and Taplin 1993), for example. However, it is possible that some, as yet unspecified, task demands of Klinger and

Dawson's experiment interacted with developmental factors to affect task performance.

Another methodological difference concerned the nature of the test question. In Klinger and Dawson's (2001) prototype condition, participants were expected to select the best category member. This requirement was either implicit or made explicit if the participant sought clarification. It is possible that such a requirement presented greater difficulty for the autism group. Both items of each pair presented in the test phases looked as if they belonged to the same target category and so there was no clear right or wrong answer. This created ambiguity. In addition, no explicit or implicit rule was provided to aid the selection of the best item. This type of ambiguity was absent in the tasks that the autism groups were successful at. In Klinger and Dawson's rule-based conditions, there was only one correct answer: only one item of each test pair was a member of the target category. Similarly, in our study (Molesworth et al. 2005) there was only a single correct answer; either the test item had been seen before or not. Furthermore, in these tasks, either implicit or explicit rules were provided. For example, in our studies, participants were taught how to indicate recognition responses. As discussed earlier, in Klinger and Dawson's rule-based conditions participants learnt the correct classification rule.

This account is plausible given that individuals with autism are known to have difficulty with the pragmatic aspects of language use (Baron-Cohen 1988; Dennis et al. 2001; Eales 1993; Surian et al. 1996; Tager-Flusberg 1981). Using context to decode ambiguity in either spoken or written language is a particular difficulty. For example, LFA children were likely to use context to determine the correct pronunciation of the word *bow* in such sentences as: 'He had a pink bow' and 'He made a deep bow' (Frith and Snowling 1983). The effect has been noted also in HFA and Asperger syndrome groups (Happé 1997; Jolliffe and Baron-Cohen 1999). Jolliffe and Baron-Cohen (1999) report, also, that HFA and Asperger syndrome groups had difficulty using context to disambiguate the meaning of spoken sentences such as: 'The roar of the fans disturbed the team' where such sentences were preceded by a contextualizing sentence, for example 'The boiler house was noisy'.

Experiments 1 and 2 were designed to test the possibility that children with autism failed to show a prototype effect in Klinger and Dawson's (2001) study because of difficulty interpreting the ambiguous task requirement rather than any impairment in prototype formation per se. HFA children and comparison children completed a prototype effect test (Experiment 1) and control tasks designed to check understanding of task ambiguity and question wording (Experiment 2).

Experiment 1

The prototype effect test used here was similar to that used by Klinger and Dawson (2001) but with a few differences, designed to increase test sensitivity. Asking participants to study six categories instead of two increased the range on the dependent variable. Items that bore low family resemblance (FR) or global similarity to the prototype were added to the test phase to provide a more stringent test. To show a prototype effect, participants should choose both more prototypes and fewer low FR items than medium FR items. The wording of the test question (e.g. 'Where is the best Hov?') made the requirement to choose the best category member explicit for all participants. If HFA participants have trouble understanding the best test question, or are not using prototypical information for categorical judgments, then they should show a reduced or absent prototype effect.

Method

Participants

Two groups took part in the study: HFA children and a typically developing comparison group. Two children with HFA and one without were excluded because of a recorded history of epilepsy or ADHD. The remaining participants were matched on gender (2 girls and 16 boys per group), and globally matched on CA and VMA. The children in the autism group had been diagnosed by clinicians as having either Asperger syndrome (13) or autism (5) according to established criteria such as those specified by the DSM-IV (American Psychiatric Association 1994). Following Miller and Ozonoff's (2000) view of Asperger syndrome as high IQ-autism, children with both diagnoses were included within a single HFA group. They were recruited from special education facilities and ranged in age from 9 years and 5 months to 15 years and 8 months. Children in the comparison group were recruited from schools from Central and South East England. Their ages ranged from 9 years and 7 months to 15 years and 7 months. VMA was assessed by the British Picture Vocabulary Scale (Dunn et al. 1997). Table 1 summarizes participant characteristics.

Materials

Six categories of cartoon animal were created using the method described by Molesworth et al. (2005). Each was labeled (e.g. *Hov*) and structured around a central prototype. This possessed features (e.g. neck or nose) that were the category average in size. All category members possessed six features that varied along a dimension consisting of five equal steps from value 1 (small) to value 6 (large).

Table 1 Participant characteristics

	Autism group ($n = 18$)	Comparison group ($n = 18$)
Chronological age (years)		
M	13.13	12.88
SD	2.02	2.04
VMA (years)		
M	12.00	12.33
SD	3.47	3.28
Range	6.75–17.00	7.42–17.00
BPVS raw scores		
M	110.11	112.39
SD	24.59	22.19
Range	69–151	76–150

Note: VMA = verbal mental age. BPVS = British Picture Vocabulary Scale. VMA was derived from the BPVS. Maximum group difference: $t(34) = .38$, $p = .71$

All study stimuli were black line drawings occupying a maximum area of 9 cm by 10 cm on white 20 cm by 12.5 cm cards. Eight study items were created for each category. The features of each study item had values of either 2 or 5. These bore medium FR, an intermediate level of similarity, to their respective category prototypes (each with feature values all at 3.5).

Test items were printed in the form of a booklet for each participant. The last eight pages contained stimuli for Experiment 2. Each page illustrated a prototype, an unstudied medium FR exemplar, and a low FR exemplar from the same category. Figure 1 illustrates one page of items belonging to the Hov category.

Each booklet presented to each HFA participant was assigned to one of two counterbalancing orders: Set A or Set B. On each page of the section of the Set A booklet used in this experiment, the position of each exemplar type

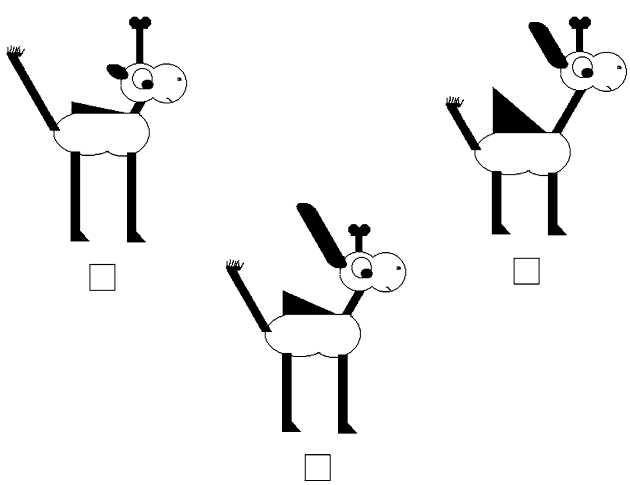

Fig. 1 Test page for the Hov category. Items from left to right are the following exemplar types: low FR, medium FR, prototype

was counterbalanced across categories and similarly with Set B booklets. However, the position–category configurations for each exemplar type were varied from Set A. For each prototype test, page order and category order were randomized. The comparison group received replicas of these booklets.

Procedure

For each participant, the sixteen study items from the pair of categories depicted on the first two pages of the test booklet were shuffled together. The first item was placed face up towards one side of the participant and named (e.g. *Hov*). The experimenter (first author) told the participant to study this and all further cards for 3 min because there would be a memory test later. Three minutes was the maximum amount of time that some of the younger participants with autism could be prompted to study the cards. The experimenter selected an item belonging to the other category, placed it towards the other side of the participant, and named it (e.g. *Mek*). From then on study cards were handed singly to the participant who was encouraged to study the card and then place it face up on the pile of cards from the same category. Any mistakes in placing the cards were corrected immediately by the experimenter. Immediately afterwards, the participant was shown the first page of the test booklet. The experimenter said, ‘Look at all these’, pointed briefly to each exemplar from left to right, and asked where the best category item was: for example, ‘Where is the best Hov?’ If the participant did not respond immediately, the question was repeated together with the comment, ‘There is no right or wrong answer, just choose the one that you think is best’. Any hesitant participants responded after a second prompt. Participants indicated selections by marking the box beneath the chosen item. They were then asked to select the best item from the other category of the study pair depicted on the second page. This study and test procedure was repeated twice more. In this way, each participant studied and was tested on all six categories, one pair at a time.

Results and Discussion

Across the six categories, each exemplar type could be chosen as best between 0 and 6 times. The total number of prototypes selected by each participant was counted and converted to a proportion out of 6. These choice proportions (CPs) were calculated for medium FR items and low FR items also. The mean CP for each exemplar type and participant group is displayed in Table 2. Support for the idea that a prototype effect shown in response to best test questions would be impaired in autism was somewhat equivocal. The data illustrated in Table 2 suggests that

Table 2 Results of the prototype effect test

Exemplar type	Autism group CP	Comparison group CP
Prototype		
M (SD)	.49 (.27)	.57 (.23)
Medium FR		
M (SD)	.28 (.16)	.30 (.18)
Low FR		
M (SD)	.23 (.23)	.13 (.17)

Note: CP = choice proportion; FR = family resemblance

both participant groups showed a prototype effect. The mean CPs increased as similarity to the prototype increased. In addition, the autism group appeared to show a weaker effect: They selected fewer prototypes and more low FR items than the comparison group. A Friedman test showed that the difference between exemplar types was significant for the comparison group: Chi-Square = 15.61, $df = 2$, $p < .001$; but that the difference between exemplar types only showed a trend towards significance for the HFA group, Chi-Square = 5.32, $df = 2$, $p = .07$. However, this apparent difference between participant groups was not supported by t-tests. No significant participant group differences were observed for either the low FR items, or the prototypes.

A full test of the original hypothesis required participants to complete control tests. If individuals with autism fail to show a prototype effect because of difficulty with ambiguity inherent in the test, then they should show similar difficulty with a control task possessing ambiguity without the requirement to abstract prototypes. Such a task, the shapes test was presented in Experiment 2, together with another control task, the numbers test, designed to assess comprehension of the test question.

Experiment 2

Participants completed both control tasks. The shapes test was designed to possess ambiguity similar to that present in Klinger and Dawson's (2001) prototype condition and Experiment 1, the current study. Participants were asked to select the best category member from an array of candidates and no rule was provided to aid with selection.

The numbers test was structurally identical to the shapes test. However, the former lacked ambiguity in that a selection rule was provided and for each question, there existed a single objectively correct answer. If difficulties lie specifically with ambiguity then impairment should be observed with the shapes test only. Furthermore, if difficulty with ambiguity is responsible for the weaker prototype effect shown by the HFA group in Experiment 1 then there should be a relationship between performances on the shapes test and the prototype effect test.

Method

Participants

The same participants from Experiment 1 took part.

Materials

The last eight pages of the booklet described in Experiment 1 formed the control tasks. Each page of the shapes test depicted six items, presented in a row, with a response box beneath each item. Within each row was a pair of canonical shapes or letters as follows: letter C and letter O, letter H and letter A, square and diamond, and circle and oval. The remaining four items of each row were hybrids representing intermediate points along a continuum of similarity between the two canonical items. These intermediates were spaced evenly across the continuum. For example, varying the size of the gap at the apex of the letter A created intermediates of the H-to-A array. The size of this gap increased by a standard measurement (2 mm) as the intermediate approximated the canonical letter H. This array was similar to that created by Hampton (1996). The remaining arrays were created specifically for the study. See Fig. 2 for an illustration of the shapes test arrays.

Each page of the numbers test presented a table of numbers as if they were school test results. The top row

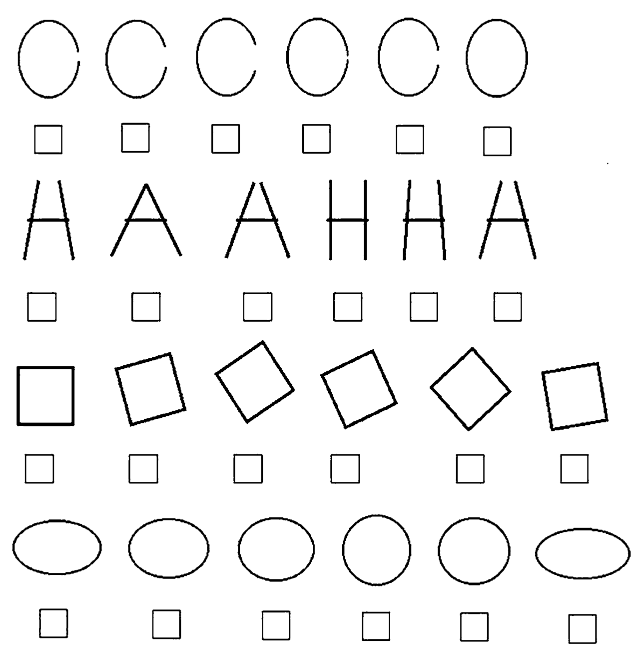

Fig. 2 Shape test stimuli. Each row of shapes or letters together with a row of response boxes was presented on a separate page

listed the subjects: English, mathematics, French, or science. The second row listed children's names; different for each school subject. The third row listed the test scores which varied across subject areas. The fourth row was left empty for the participants to place their responses.

To avoid training effects from the shapes and numbers tests, all participants completed the prototype effect test (Experiment 1) first. The position of shapes and numbers test items within each array varied randomly on each page. These random orders were held constant for each counterbalancing order, Set A or Set B. The order of arrays was randomized within each booklet and the presentation order of the control tasks was counterbalanced across each set. Comparison group participants received replica test booklets.

Procedure

For each page of the shapes test the experimenter said, 'Look at all these' and pointed briefly to each item from left to right. Then the participant was asked to point to the target canonical item, for example, 'Where is the best letter *H?*' Other targets comprised the letter *C*, the square, and the circle. The participant responded by marking a response box. At the first page of the numbers test, the participant was told that the numbers represented test marks for each of the children named, that high numbers were 'good', and low numbers were 'bad'. At each page the participant was told to look at all the numbers and asked, 'Who has the best science (mathematics, English, or French) score?' The participant responded by marking the row beneath one of the numbers.

Results and Discussion

Shapes Test

Each selection from each array of the shapes test was assigned an integer from 1 to 6. These integers reflected similarity between the selected item and the target canonical item: for example, 6 was assigned to the correct canonical item, 5 was assigned to the next most similar item, and so forth. The integers, corresponding to the items chosen from each array, were summed to give a total score (maximum = 24). This was then converted to proportions to give the proportion of shapes score (PS) for each participant. Every comparison group participant obtained the maximum PS of 1. The mean PS of the HFA group was .97 (SD = .04). The difference between participant groups was significant: $t(17) = 2.61$, $p = .02$ (equal variances not assumed).

To explore a possible relationship between PS and the developmental variables in the HFA group, CA and VMA were split on the mean PS for all participants (.99). Those in the low PS group, scoring below the mean, had a lower average CA ($M = 11.89$ years) and lower average VMA ($M = 10.61$) than the high PS group that scored above the mean: CA: $M = 13.76$ years; VMA: $M = 12.70$ years. The difference in CA showed a trend towards significance: $t(16) = 2.00$, $p = .06$. The difference in VMA was not significant although, with a sample size of six, power was low. There was no evidence that PS scores were split on formal diagnosis: Amongst those that failed to show a prototype effect, four had a diagnosis of Asperger syndrome and two had a diagnosis of autism. A Fisher's Exact Test revealed no statistically significant association between diagnostic category and prototype effect performance.

Numbers Test

As with the shapes test, each selection from each table of the numbers test was scored separately and assigned an integer from 1 (for the lowest test score) to 6 (the highest test score). The integers were summed to give a total (maximum = 24) and converted to proportions to give the proportion of numbers score (PN) for each participant. The mean PN of the HFA group and comparison group was .96 (SD = .11) and 1 (SD = .01) respectively. An independent samples *t*-test revealed no statistically significant difference between participant groups.

The finding of group differences on the shapes test is in keeping with the prediction made earlier that HFA participants would have trouble with this task if they had difficulty understanding the ambiguity. However, this conclusion applies only to one third of HFA participants tested here because the remainder performed at ceiling on the test. Performances at ceiling on the numbers test indicated that both participant groups understood the best test question used with an unambiguous task.

Relationship Between Prototype Effect Test and Shapes Test

Although there was no statistically significant difference between groups on the prototype effect test, the HFA group appeared to show a somewhat weaker effect. To see if there was any relationship between the shapes test scores and the prototype effect test scores, participants were split into three groups: Six HFA participants who scored below the mean (HFA low scorers), twelve HFA participants who scored above the mean (HFA high scorers) and eighteen comparison participants that also scored above the mean. Figure 3 illustrates the CP means of each exemplar type for each of these groups. This shows almost identical prototype effects for the HFA high scorers and the comparison group. In contrast, means obtained by the HFA low scorers do not

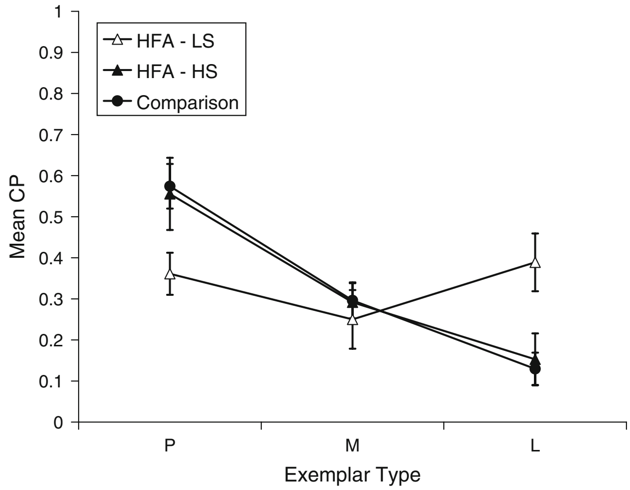

Fig. 3 Mean choice proportion (CP) for each participant group and exemplar type. P = prototype, M = medium FR, L = low FR. HFA-LS = HFA participants scoring below the mean on the shapes test ($n = 12$). HFA-HS = HFA participants scoring above the mean on the shapes test ($n = 6$). Comparison group: $n = 18$. Error bars represent standard error of the mean

form a prototype effect as shown by the relatively high CP mean for low FR items and the relatively low CP mean for prototypes. Consistent with this observation, Friedman tests revealed a significant difference between exemplar types for the HFA high scorers: Chi-Square = 6.89, $df = 2$, $p = .03$, but no significant difference for the low scorers.

To examine participant group differences directly, the presentation order of the control tasks was included in the following analysis of choice proportion scores for prototypes: A 3 (group) × 2 (order) ANOVA, where group comprised HFA high scorers, HFA low scorers and the comparison group. Levene's test of equality of error variance was significant at $F(5,30) = 4.87$, $p = .002$. Neither of the main effects or the interaction was significant. Games-Howell post hoc tests revealed significant differences between the HFA low scorers and the comparison group: mean difference = .21, $p = .03$. No other differences were significant.

The CP scores for low FR exemplars were analyzed by a 3 (group) × 2 (order) ANOVA. The main effect of group was significant: $F(2,30) = 4.30$, $p = .02$. Neither the main effect of order nor the interaction was significant. Games-Howell post hoc tests revealed significant differences between the HFA low scorers and the comparison group: mean difference = .26, $p = .03$. None of the other differences were statistically significant.

In keeping with our prediction there did seem to be an association between performance on the shapes test and performance on the prototype effect test. HFA participants that failed to perform at ceiling on the shapes test also failed to show a prototype effect.

General Discussion

The majority of HFA participants showed a clear prototype effect using a similar categorization test to that used by Klinger and Dawson (2001). Thus a strong version of our hypothesis, that the HFA group as a whole, would be affected by task ambiguity and fail to show an effect was unsupported. Additionally, this finding adds weight to our earlier research demonstrating that HFA children show intact prototype effects (Molesworth et al. 2005). It appears that Klinger and Dawson's proposal that individuals with autism show diminished prototype effects does not extend to individuals at the high functioning end of the spectrum.

Another key finding was that HFA performance on the prototype effect task was mixed. One third of HFA participants failed to show any prototype effect, the remainder did show an effect, identical to that shown by the comparison group. We speculate that this mixed performance reflected differences within the HFA group.

One form of heterogeneity was developmental in nature. The variables of CA and VMA covered considerable ranges of seven years and ten years respectively for both participant groups. It is possible that the relationship between these variables and task performance differed between participant groups. This possibility arises because autism is characterized by pronounced peaks and troughs in abilities within cognitive, linguistic and social domains (Burack et al. 2004; Jarrold and Brock 2004). For example, nonverbal abilities tend to be higher than verbal abilities (Joseph et al. 2002). Additionally, within the verbal domain the use of a verbal label to identify objects, as measured by the BPVS, reflects a peak ability (Mottron 2004). If successful prototype effect performance required the late-maturing of a 'trough' ability within the HFA group, then higher thresholds for CA and VMA would be required to demonstrate prototype effects. A similar explanation appears to hold for another task, the false belief test. Participants with autism require higher VMA thresholds than comparison children before succeeding on this task (Happé 1995). Perhaps this account could also explain the emergence of the sub-group within the HFA group of the present study. The LFA participants in Klinger and Dawson's study all failed to show a prototype effect and had a lower VMA than participants in the present study. Additionally, in the latter study there was a trend for individuals who failed the prototype effect test to have lower VMA and CA. This observation was not supported statistically, although small sample sizes meant that the statistical tests were underpowered.

Language comprehension is a plausible trough ability. Those individuals with autism that failed to show a prototype effect in the present study and in Klinger and

Dawson's study may not have understood the instructions given by the experimenters in the same manner as the comparison group, quite apart from issues related to task ambiguity. Participants from both these studies were matched on the BPVS or an equivalent. If the BPVS actually measures a peak ability as suggested earlier, then the autism groups would have had lower general language comprehension than the comparison groups.

The main difficulty with this explanation concerns the use of control conditions by Klinger and Dawson's study and the present one; the rule-based categorization test and the numbers test. On both these tests, autism performance matched that of comparison children and so a difficulty with language comprehension was not implicated. The question wording used on these tests was almost identical to that used for the prototype effect tests which proved problematic. Success on these control tasks therefore implies that language comprehension was not responsible for failure to show prototype effects. The explanation could be salvaged, however, if it were demonstrated that the sensitivity of these control tests was questionable. This is possible because performances on these were either at or close to ceiling.

Other forms of heterogeneity were not developmental; formal diagnosis, for example. As mentioned earlier, individuals with Asperger syndrome and autism were included within the same experimental group. This variable, however, did not appear to delineate the HFA subgroup. Both individuals with Asperger syndrome and autism were represented within it.

Quite apart from developmental variables and diagnosis, considerable heterogeneity has long been recognized as characteristic of the autism spectrum. For example, the expression of social impairments varies. Wing and Gould (1979)'s epidemiological study of autism found that some children with autism could be described as withdrawn and aloof, others as socially active but inappropriate and others showed unusual passivity in relation to social situations. General recognition of this heterogeneity has led some researchers to speculate that inconsistent findings between studies can be attributable to the fact that subgroups exist that vary in their likelihood to show diminished performance on the task of interest. For example, Ropar and Mitchell (2001) failed to replicate Happé's (1996) finding that individuals with autism do not succumb to visual illusions. They suggested that subgroups existed within the autism population that varied in susceptibility to these illusions. In a similar vein, sub-groups may exist that vary in their susceptibility to prototype effects. Unfortunately the present study provided no non-developmental data on what the defining characteristics of this sub-group might be.

Another issue arising from the findings of the present study is the implications of the association between the shapes test and the prototype effect test. Those HFA individuals that did not perform at ceiling on the shapes test were those that failed to show a prototype effect. In principle it is possible that lower performance on these tests can be attributable to different causes. It is more parsimonious to assume however that the two tasks share common features that are problematic for the low scoring HFA group. The tasks were designed to have ambiguity as a common feature, and the findings are consistent with a weaker version of our original hypothesis: that a minority of HFA participants failed to show a prototype effect because of difficulty with ambiguity inherent in the task.

A full account of the findings of the present study would need to explain the interaction between the HFA group heterogeneity and the association between performances on the prototype effect test and the shapes test. If, for example, CA and VMA are the critical defining features of the subgroup and if ambiguity is a critical feature held in common between the tasks then these developmental variables must impact upon the appreciation of ambiguity. HFA individuals might need to be of a certain age and VMA before they possess the ability to resolve the ambiguity necessary to succeed on these tasks. At present, there is insufficient evidence to support such an account. Those studies that have researched pragmatic function, a useful tool for resolving ambiguity within language, tend to report strongly diminished performance in the pragmatic use and understanding of language within autism. This difficulty can not be attributable to problems with general language use (e.g. Dennis et al. 2001; Eales 1993; Surian et al. 1996).

One limitation of the current study is that performance on the prototype effect was heterogeneous and there was no evidence, supported by statistical significance, regarding the defining characteristics of the HFA subgroup. This perhaps reflects the choice of experimental design. The current study used the design that Jarrold and Brock (2004) argue is the predominant one in autism research; one of factorial comparison with a focus on group differences. This design is ill-suited to uncovering the variables affecting within-group differences such as those obtained in the current study. Instead, it is likely that one that analyzes patterns of association between experimental variables and focuses on individual differences will yield more information on the variables governing prototype effect performance in autism. The preceding discussion has identified some candidate variables that could affect performance on the prototype effect. These include the developmental variables CA and VMA and measures of language comprehension, pragmatic function and adaptive social functioning. Future research on this topic could use regression or ANCOVA techniques as recommended by Jarrold and Brock to uncover the relationship of these variables to task performance.

Acknowledgments Thanks are due to the participants and their teachers for their invaluable assistance. Molesworth would also like to thank Claire Simmons for her comments on the data and on an earlier draft of this paper. The data presented here was submitted by Catherine Molesworth as part of her Ph.D. thesis (Department of Psychology, City University).

References

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

Baron-Cohen, S. (1988). Social and pragmatic deficits in autism: Cognitive or affective? *Journal of Autism and Developmental Disorders, 18*, 379–402.

Bennetto, L., Pennington, B. F., & Rogers, S. J. (1996). Intact and impaired memory functions in autism. *Child Development, 67*, 1816–1835.

Berger, H. J., Van Spaendonck, K. P., Horstink, M. W., & Buytenhuijs, E. L., et al. (1993). Cognitive shifting as a predictor of progress in social understanding in high-functioning adolescents with autism: A prospective study. *Journal of Autism and Developmental Disorders, 23*, 341–359.

Bowler, D. M., Gardiner, J. M., Grice, S., & Saavalainen, P. (2000). Memory illusions: False recall and recognition in adults with Asperger's syndrome. *Journal of Abnormal Psychology, 109*, 663–672.

Bowler, D. M., Matthews, N. J., & Gardiner, J. M. (1997). Asperger's syndrome and memory: Similarity to autism but not amnesia. *Neuropsychologia, 35*, 65–70.

Burack, J. A., Iarocci, G., Flanagan, T. D., & Bowler, D. M. (2004). On mosaics and melting pots: Conceptual considerations of comparison and matching strategies. *Journal of Autism and Developmental Disorders, 34*, 65–73.

Cabeza, R., Bruce, V., Kato, T., & Oda, M. (1999). The prototype effect in face recognition: Extension and limits. *Memory and Cognition, 27*, 139–151.

Dennis, M., Lazenby, A. L., & Lockyer, L. (2001). Inferential language in high-function children with autism. *Journal of Autism and Developmental Disorders, 31*, 47–54.

Dunn L. M., Dunn L. M., Whetton C., & Burley, J. (1997). *The British picture vocabulary scale*. London: NFER-Nelson.

Eales, M. J. (1993). Pragmatic impairments in adults with childhood diagnoses of autism or developmental receptive language disorder. *Journal of Autism and Developmental Disorders, 23*, 593–617.

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

Hampton, J. A. (1996). Conjunctions of visually based categories: Overextension and compensation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 378–396.

Hampton, J. A. (1997). Psychological representation of concepts. In M. A. Conway (Ed.), *Cognitive models of memory studies in cognition* (pp. 81–110). Cambridge, MA: The MIT Press.

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

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

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

Hayes, B. K., & Conway, R. N. (2000). Concept acquisition in children with mild intellectual disability: Factors affecting the abstraction of prototypical information. *Journal of Intellectual and Developmental Disability, 25*, 217–234.

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

Hermelin, B., & O'Connor, N. (1970). *Psychological experiments with autistic children*. Oxford: Plenum Press.

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

Jarrold, C., & Brock, J. (2004). To match or not to match? Methodological issues in autism-related research. *Journal of Autism and Developmental Disorders, 34*, 81–86.

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

Joseph, R. M., Tager-Flusberg, H., & Lord, C. (2002). Cognitive profiles and social-communicative functioning in children with autism spectrum disorder. *Journal of Child Psychology and Psychiatry, 43*, 807–821.

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* (pp. 119–136). New York: Plenum Press.

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

Metcalfe, J., & Fisher, R. P. (1986). The relation between recognition memory and classification learning. *Memory and Cognition, 14*, 164–173.

Miller, J. N., & Ozonoff, S. (2000). The external validity of Asperger disorder: Lack of evidence from the domain of neuropsychology. *Journal of Abnormal Psychology, 109*, 227–238.

Minshew, N. J., Goldstein, G., Muenz, L. R., & Payton, J. B. (1992). Neuropsychological functioning nonmentally retarded autistic individuals. *Journal of Clinical and Experimental Neuropsychology, 14*, 749–761.

Molesworth, C. J., Bowler, D. M., & Hampton, J. A. (2005). The prototype effect in recognition memory: Intact in autism?. *Journal of Child Psychology and Psychiatry, 46*, 661–672.

Mottron, L. (2004). Matching Strategies in Cognitive Research with Individuals with High-Functioning Autism: Current Practices, Instrument Biases, and Recommendations. *Journal of Autism and Developmental Disorders, 34*, 19–27.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge: MIT Press.

Omohundro, J. (1981). Recognition vs. classification of ill-defined category exemplars. *Memory and Cognition, 9*, 324–331.

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.

Plaisted, K. (2001). Reduced generalization in autism: An alternative to weak central coherence. In J. A. Burack, T. Charman, et al. (Eds.), *The development of autism: Perspectives from theory and research* (pp. 149–169). Mahwah, NJ: Lawrence Erlbaum Associates.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77*, 353–363.

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

Rosch, E. (1975a). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*, 192–233.

Rosch, E. (1975b). The nature of mental codes for color categories. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 303–322.

Rosch, E., Simpson, C., & Miller, R. S. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance, 2*, 491–502.

Solso, R. L., & McCarthy, J. E. (1981). Prototype formation of faces: A case of pseudo-memory. *British Journal of Psychology, 72*, 499–503.

Surian, L., Baron-Cohen, S., & Van der Lely, H. (1996). Are children with autism deaf to Gricean maxims? *Cognitive Neuropsychiatry, 1*, 55–71.

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

Tager-Flusberg, H. (1981). On the nature of linguistic functioning in early infantile autism. *Journal of Autism and Developmental Disorders, 11*, 45–56.

Tager-Flusberg, H. (1985a). Basic level and superordinate level categorization by autistic, mentally retarded, and normal children. *Journal of Experimental Child Psychology, 40*, 450–469.

Tager-Flusberg, H. (1985b). The conceptual basis for referential word meaning in children with autism. *Child Development, 56*, 1167–1178.

Tager-Flusberg, H. (1991). Semantic processing in the free recall of autistic children: Further evidence for a cognitive deficit. *British Journal of Developmental Psychology, 9*, 417–430.

Ungerer, J. A., & Sigman, M. D. (1987). Categorization skills and receptive language development in autistic children. *Journal of Autism and Developmental Disorders, 17*, 3–16.

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders, 9*, 11–29.

Younger, B. (1990). Infant categorization: Memory for category-level and specific item information. *Journal of Experimental Child Psychology, 50*, 131–155.