This article may not exactly replicate the final version published in the APA journal. It is not the copy of the record. DOI - 10.1037/a0038204

Word count: 7471

Children with Autism do not Show Sequence Effects with Auditory Stimuli

Catherine Molesworth, Coralie Chevallier, and Francesca Happé

King's College London, London, UK

James A. Hampton

City University London, London, UK

Author note

Catherine Molesworth, Coralie Chevallier, and Francesca Happé, MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, UK. James Hampton, Department of Psychology, City University London, London, UK.

Catherine Molesworth is now at London South Bank University, UK. Coralie Chevallier is now at the Laboratoire des Neurosciences Cognitives, INSERM U960, Ecole Normale Supérieure, Paris, France.

Acknowledgments. CM and FH were supported by the Economic and Social Research Council [grant code PTA-026-27-2226]. CC was supported by the Economic and Social Research Council [grant code RES-000-22-3136]. CC was also supported, in part, by ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL*. We would like to thank Neil Stewart for supplying the 42 stimulus presentation orders and the children and staff who assisted with the study.

Correspondence should be addressed to Catherine Molesworth, Department of Psychology, London South Bank University.

E-mail: c.j.molesworth@gmail.com

Categorization decisions that reflect constantly changing memory representations may be an important adaptive response to dynamic environments. We assessed one such influence from memory, sequence effects, on categorization decisions made by individuals with autism. A model of categorization (i.e. Memory and Contrast model, Stewart, Brown, & Chater, 2002) assumes that contextual influences in the form of sequence effects drive categorization performance in individuals with typical development. Difficulties with contextual processing in autism, described by the weak central coherence account (Frith, 1989; Frith & Happé, 1994) imply reduced sequence effects for this participant group. The experiment reported here tested this implication. High functioning children and adolescents with autism (aged 10 to 15 years), matched on age and IQ with typically developing children, completed a test that measures sequence effects (i.e. category contrast effect task, Stewart et al., 2002) using auditory tones. Participants also completed a pitch discrimination task to measure any potential confound arising from possible enhanced discrimination sensitivity within the ASD group. The typically developing group alone demonstrated a category contrast effect. The data suggest that this finding cannot be attributed readily to participant group differences in discrimination sensitivity, perseveration, difficulties on the associated binary categorization task, or greater reliance upon long term memory. We discuss the broad methodological implication that comparison between autism and control group responses to sequential perceptual stimuli may be confounded by the influence of preceding trials. We also discuss implications for the weak central coherence account and models of typical cognition.

Key words: Autism, categorization, pitch discrimination, sequence effects, weak central coherence

Autism spectrum disorders (ASDs) are neurodevelopmental conditions diagnosed on the basis of behaviors such as impairments in communication and reciprocal social interaction. Early researchers into autism suggested that difficulties with conceptual representation and categorization were also central to the condition (Hermelin & O'Connor, 1970; Rimland, 1964; Scheerer, Rothmann, & Goldstein, 1945). Some researchers have suggested that such difficulties may underpin or exacerbate difficulties processing social information (Gastgeb, Strauss, & Minshew, 2006; Klinger & Dawson, 1995). For example, a father struggled to teach his son with autism to stay away from strangers (Klinger & Dawson, 1995). The son could not readily grasp the meaning of the concept and kept asking for a set of *if then* rules to define the category.

Despite these theoretical and anecdotal suggestions of a categorization impairment, the application of traditional models of categorization (e.g. Murphy, 2002) has demonstrated that categorization processes are broadly typical for individuals with autism who do not have intellectual disability. For example, once familiar with a data set that permits both exemplar-based and rule-based strategies, participants with autism are as likely to use rules to categorize as typical participants (Soulières, Mottron, Giguère, & Larochelle, 2011). Individuals with autism in general appear equally subject to the influence of prototypes as control participants (e.g. Gastgeb et al., 2006; Molesworth, Bowler, & Hampton, 2005; Vladusich, Lafe, Kim, Tager-Flusberg, & Grossberg, 2010). One study that modeled categorization performance with a quantitative exemplar model, the general context model (GCM, Nosofsky, 1986) reported no differences between an ASD group and control group. Once trained to criterion both participant groups categorized test phase items with similar accuracy and discriminated similarly between memory traces for exemplars (Bott, Brock, Brockdorff, Boucher, & Lamberts, 2006).

As the above findings demonstrate, participants with autism can categorize stimuli into binary categories as successfully as participants with typical development (TD). The findings also show that traditional models of categorization do not describe categorization processes that are problematic for individuals with autism. Such models assume that categorization is made on the basis of absolute magnitude information available from long term memory and do not represent influences from short term memory. In contrast, relative judgment models such as the memory and contrast (MAC) model (Stewart & Brown, 2004; Stewart et al., 2002), assume that absolute magnitude information is either unused or unavailable from long term memory. The MAC model represents only short term contextual influences, such as feedback and stimulus percepts from previous trials. As we argue later, this type of model may better describe categorization processes that are problematic in autism.

Evidence that absolute magnitude information is not used comes from estimates of tone loudness, for example. These reflect the perceived magnitude of previous tones (i.e. sequence effects) rather than absolute magnitude information (Baird, Green, & Luce, 1980). To illustrate the influence of sequence effects on categorization, Stewart et al. (2002) designed the category contrast effect task. This paradigm uses stimuli comprising ten auditory tones (Tone 1 to Tone 10) spaced along a unidimensional continuum of pitch. Each tone represents a constant increase in perceived pitch that is associated with an increase in sound frequency by a constant ratio (Krumhansl, 2000; Shepard, 1982). Tones were presented in sequence and with feedback participants learnt to categorize them into two categories; low pitch (the five lowest tones) versus high pitch (the five highest tones). The key variable was response accuracy to the two tones closest to category boundary (Tone 5 and Tone 6). Each boundary tone was preceded by a distant tone (Tone 1 or Tone 10).

These distant tones were either from the same category as the boundary tone or from the contrast category. See Figure 1 for an illustration of the ten tones. The results demonstrated a *category contrast effect*: Response accuracy to boundary tones was greater following a distant tone from the contrast category, than following a distant tone from the same category.

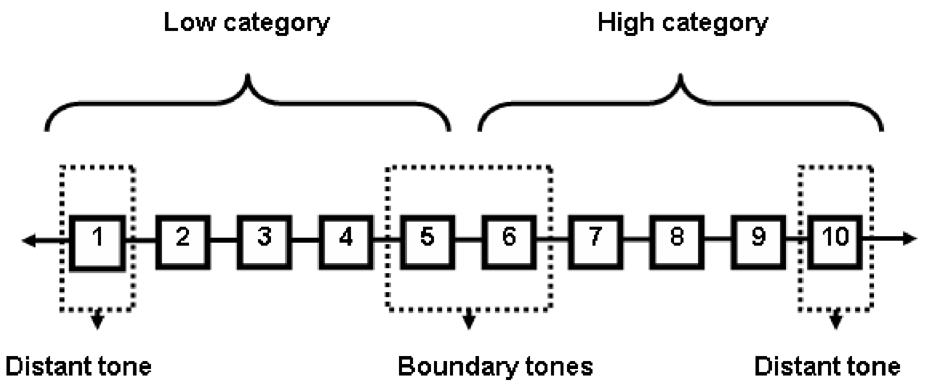

Figure 1. Illustration of ten tones used to demonstrate a category contrast effect (Stewart et al., 2002).

Explanations for the category contrast effect all assume a *recency bias*. This refers to an inherent tendency to weight immediately preceding information most heavily (M. Jones, Love, & Maddox, 2006). Stewart and colleagues focus on decisional processes (Stewart & Brown, 2004; Stewart et al., 2002). For example, Stewart et al. (2002) suggest that participants adopt heuristic decision making strategies when encountering sequential categorization tasks, none of which require access to absolute magnitude information from long term memory. If for example, the participant learns that one tone is “high”, and the next tone is even higher, they know that tone will be high too. If the next tone is lower than the high tone, their decision will depend on the magnitude of perceptual difference between the current and preceding tone. For small changes, they are likely to repeat the original response, but larger changes induce a switch in category label. These strategies produce contrast effects in sequential stimuli.

Stewart and colleagues maintain that sequence effects are the sole driver of categorization with simple perceptual stimuli (Stewart & Brown, 2004). They accept, however, that information from long term memory may also be used (Stewart & Matthews, 2009). There is evidence, for example, that participants rely more upon long term memory when a minority of feedback is unreliable as in probabilistic categorization tasks (Craig, Lewandowsky, & Little, 2011) or when feedback is ceased part way through the categorization task (Nosofsky & Little, 2010). However, Stewart and Matthews argue that such long term memory representations comprise *difference* information (presumably derived from sequence effects) with no absolute magnitude representations. This view may not apply to individuals with autism, because although they can perform binary categorization, there is evidence that they have difficulty with contextual processing, defined broadly. This difficulty may include the type of processing involved in producing the category contrast effect. Context in this case would comprise immediately preceding information such as stimulus percepts and feedback.

Some of the contextual processing difficulties shown by individuals with autism involve visual stimuli with a target item presented simultaneously with context. For example, adults with autism were less efficient at locating objects that were conceptually incongruent with surrounding context in depictions of common scenes (e.g. a butterfly in a winter snow scene, Jolliffe & Baron-Cohen, 2001). Difficulties have also been observed with verbal stimuli and context unfolding over time such as during homograph tasks (Frith & Snowling, 1983; Happé, 1997). These require the participant to read aloud either frequent or infrequent pronunciations of homographs embedded within a sentence that provides context (e.g. “In her eye / dress there was a big tear”, Happé, 1997). Participants with autism show a tendency to produce contextually inappropriate frequent pronunciations of homographs. These findings are captured by a general explanation of cognitive processing in autism, the weak central coherence account (Frith, 1989; Happé, 1994a). Frith (1989) coined the term “central coherence” to refer to the natural human tendency to “draw together diverse information to construct higher-level meaning in context” (Frith & Happé, 1994, p.121). Frith argued that this tendency is weakened in autism.

The category contrast effect task, like the homograph task, entails context unfolding over time. If individuals with autism are subject to less contextual influence, as suggested by the weak central coherence account, they should show a reduced category contrast effect. To assess this, we presented a test of sequence effects, the category contrast effect task (Stewart et al., 2002), to children with autism and a control group of children with TD.

An alternative explanation for any reduced category contrast effect could be that it arises from a tendency for participants with autism to weight information from long term memory at the expense of information from short term memory, rather than contextual processing difficulties per se. This account is consistent also with findings that individuals with autism favor the most frequent homograph pronunciations over the most contextually appropriate ones. To assess this we present a model for each participant group that identifies the respective contributions from these different memory stores.

We also measured pitch discrimination ability because this could influence the size of the category contrast effect. For example, enhanced ability to discriminate between boundary tones may make a rule defining category membership more readily available. Use of such a rule would preclude any need to depend upon sequence effects and therefore reduce the size of the category contrast effect.

There is evidence that participants with autism may have superior discrimination ability. One study has shown enhanced pitch discrimination at group level (Bonnel et al., 2003). Two other studies reported enhanced discrimination within sub-groups of participants with autism; approximately 10% of the ASD samples achieved performance that measured 2 *SD* or higher above the TD group mean on pitch discrimination tasks (Heaton, Williams, Cummins, & Happé, 2008; C. Jones et al., 2009). Assessing this ability is therefore necessary to interpret the size of the category contrast effect.

Method

Participants

Thirty-two male adolescents and children took part in the study comprising 15 ASD participants and 17 TD participants. Groups were matched on chronological age and FSIQ as measured by the Wechsler Abbreviated Scales of Intelligence, two test form (Wechsler, 1999). See Table 1 for a summary of participant characteristics.

Table 1

Participant characteristics: mean, (standard deviation) and *range*

	ASD (*n*=15)			TD (*n*=17)		
Age (years)	13.14	(1.41)	*9.92-14.92*	13.16	(1.56)	*9.83-14.92*
FSIQ	99.27	(15.70)	*61-119*	103.12	(13.00)	*68-118*
ADOS	12.53	(3.81)	*7-18*		N/A	

Both participant groups were recruited from schools located in London, and South East or East England. Participants were screened for hearing impairment by consulting teachers and school records. One potential participant with an ASD was excluded under this criterion. Children in the ASD group were recruited from specialist schools or units. All had been diagnosed by independent clinicians as having an ASD using standard criteria, such as those specified by the DSM-IV (American Psychiatric Association, 1994). Each had received a formal diagnosis of either Asperger syndrome (4), or autism (11), verified by checking school records. In the UK, entry into a special school or unit for children with autism is usually dependent upon local authority funding and is accompanied by a formal diagnosis made by a registered clinician, which is documented in medical records. We used the Autism Diagnostic Observation Schedule (ADOS, Lord et al., 2000) as an additional screening instrument for an ASD. Only those gaining a total score (a sum of Communication and Reciprocal Social Interaction subscale scores) at or above the threshold for an ASD (7 points) were included in the study. Eleven participants met or exceeded the more stringent threshold for autism (11). The local ethics committee of King's College London approved the procedure for the experiment.

Design

Participants completed two tasks; a category contrast effect task followed by a discrimination task. The order of these tasks was not counterbalanced to prevent familiarization with the stimulus set before commencing the category contrast effect task. Such familiarization could reduce the importance of sequential context, the determinant of the contrast effect.

To reduce fatigue, the tasks were presented on different days with an inter-task interval ranging from 1 to 7 days (*M* = 3 days, ASD group) or 1 to 9 days (*M* = 4 days,TD group). To compensate for variation in recency to stimulus exposure, participants completed a block (20 trials) selected randomly from the category contrast effect task immediately before commencing the discrimination task.

Category contrast effect. Following Stewart et al. (2002), Experiment 1, tones were divided into two categories. The five lowest frequency tones comprised the low category and the five highest frequency tones comprised the high category (see Figure 1). Tones were presented for categorization in sequence and the category contrast effect was dependent upon responses to four critical tone pairs. Two pairs each comprised tones from the same category; low tones (1 then 5), and high tones

(10 then 6). Two pairs each comprised tones from different categories; Tone 1 then Tone 6, and Tone 10 then Tone 5. Each of the four critical pairs occurred once in each block of 20 trials. Each block comprised a sequence selected randomly from a set of 42 sequences devised by Stewart et al. (2002). These sequences were designed to appear random but each was structured as follows: Each critical pair was randomly assigned to the 4th and 5th, 9th and 10th, 14th and 15th, and 19th and 20th trials of each block. A practice categorization task preceded the category contrast effect task.

Discrimination task. Participants completed a practice discrimination task, and the main discrimination task in succession. The design of both tasks was as follows: Each one of six stimuli was replicated a further three times to form a block of trials; 12 pairs divided equally into two categories; *Different* versus *Same*. Each category contained two instances of each tone. The computer played the tones in pairs of low distant tones, boundary (medium pitch) tones, and high distant tones. Within each pair, a standard tone was followed by a comparison tone. Within the Different category, the identity of the standard was counterbalanced across each pair; for example, for boundary tone pairs, Tone 5 was the standard in one pair, and Tone 6 was the standard in the other. Pairs in the Same category each comprised two instances of the same tone (e.g. Tone 1 then Tone 1, Tone 2 then Tone 2).

Stimuli and Apparatus

The computer presented all visual displays in yellow against a blue background.

Category Contrast Effect Practice Stimuli. Practice stimuli comprised 10 line drawings of shapes; circle, square, diamond, triangle, and pentagon. Each shape was represented twice; once as a *small* shape occupying no more than 55 mm^2 of screen space, and once as a *big* shape occupying no more than 250 mm^2 of screen space.

Category Contrast Effect Stimuli. Stimuli were ten sine wave tones generated and saved using auditory digital editing software (Audacity, Version 1.2.6). Tones differed along a single dimension, pitch, and each tone was 2% higher than the nearest lower tone. Thus, tones were designed to be equidistant along the continuum of perceived pitch as in Stewart et al. (2002). The first, Tone 1, had a frequency of 550.00Hz and the last, Tone 10, had a frequency of 657.30Hz. The frequency interval of 2% was selected in preference to the 1% interval selected by Stewart et al. to reduce potential floor effects in discrimination task data. Click artifacts were removed using the method described by Stewart and Brown (2004). A laptop computer played the tones through Sennheiser PXC250 headphones with active noise compensation (attenuation of frequencies below 1000 Hz by 15 db). Stimuli were played at 62-68 db throughout both tasks as measured by a mini sound level meter (Class IEC 651 Type II, A frequency rating) placed touching both headphone earpieces. The task was programmed with E-prime 2.0 software.

Discrimination Practice and Main Task Stimuli. Stimuli comprised a pair of tones from each distant end of the category continuum, Tones 1 and 2, Tones 9 and 10, and the boundary pair, Tones 5 and 6.

Procedure

Category Contrast Effect Practice Task. Participants were tested singly in a quiet room on school premises. They were shown a picture of a screen event, a circle flanked by the labels *Small* and *Big* and told that the computer would show more shapes like that, one presented after the other. They were told to choose if the shape was small or big and “to make their answer” to the computer by pressing one of two computer keys. The key associated with each choice was located directly below the corresponding label on screen. Then participants were told that the correct answer would remain on the computer screen and the wrong answer would disappear. Participants were told that if they took too long the computer would present the next shape without waiting for an answer. The final instructions were to answer as quickly and accurately as possible and participants were invited to ask any questions about the procedure. These were the final instructions for each of the remaining tasks also.

Each participant completed one block of 10 trials presented in random order. Each trial began with the presentation of a shape, flanked by labels to the left and right, presented for 500ms on the computer screen. From stimulus onset, participants could respond by pressing the *F* key to make a response denoted by the left hand label, or the *L* key to make a response denoted by the right hand label. Colored labels covered these response keys. The association between key and corresponding category label was counterbalanced across participants. The timing of screen events was identical to that used by Stewart et al. (2002). Either 2,000 ms after stimulus onset or after the participant responded, whichever was the sooner, the stimulus and wrong label disappeared and the correct label remained on the screen for a further 1000 ms. The next trial began after an additional delay of 500 ms. The computer displayed the sum of correct trials and average reaction time at the end of the block. One participant with autism scored 1/10 correct but after further instruction, repeated the block of trials without error. None of the remaining participants scored less than 7/10 on the practice task.

Category Contrast Effect Task. Participants completed the category contrast effect task immediately after the practice task. Instructions to participants and the procedure were identical, except as follows: Stimuli were auditory tones played over headphones. Participants were told that they would "hear some computer beeps one after the other". They were shown a picture of a screen event, a question mark flanked by the category labels, *low* versus *high*. These category labels were not used by Stewart et al. (2002) but were used in the present study as memory aids. In addition, participants were told that at first they would not "know all the right answers". They were advised to attend to the correct answers and so learn which category each beep belonged to. Participants were familiarized with the headphones prior to starting the task. One participant with autism declined to wear them, so tones were played through computer loudspeakers during the categorization and discrimination tasks for this participant. Data from this participant were not outliers and so were included in the analyses below.

Each participant completed 6 blocks of 20 trials. After every two blocks, the computer offered a break and displayed the sum of accurate responses and average reaction times across both blocks to maintain motivation. Trials in each block were ordered in a sequence selected with replacement from the set of 42 pseudorandom sequences described in the design.

Discrimination Practice Task. Participants were shown a picture of a screen event, a question-mark flanked by the labels *Same* and *Different* and told that when a similar picture appeared on the computer screen they would hear two beeps, one after the other. They were told to choose if the beeps were the same or different and "to make their answer" to the computer by pressing the *F* key or *L* key. The experimenter indicated the key associated with each choice, located directly below the corresponding label on screen. Participants were told that tones were different if one was higher or lower in pitch than the other. Then participants were instructed that after their response, the label associated with the correct answer would stay on the screen and the wrong answer would disappear. Participants were told that the computer would wait for their answer. One trial comprised one pair of tones. Two blocks, each of 12 trials described in the design (6 Different trials and 6 Same trials), were presented to each participant. Trials within each block were presented in random order without replacement. Each trial started with the standard tone, played for 500ms. Then 1000 ms after standard tone offset the comparison tone played for 500ms. From the onset of the comparison tone, participants could respond by pressing the *F* key or *L* key. The association of computer key with corresponding category label (Different vs. Same) was counterbalanced across participants. Immediately after the participant's response, the wrong label disappeared and the correct label remained on screen to provide feedback for a further 1000ms. The next trial began after an additional delay of 500ms.

Discrimination Task. Participants were told that the next "game" was just like the previous one, but that the computer would no longer show if the answer was correct or not. They were told that half the pairs of beeps were the same and that half were different and that pairs of each type would be randomly intermixed.

Participants completed six blocks, each of 12 trials. The computer offered a break and displayed the sum of accurate responses and the average reaction time after completion of three blocks. The computer again displayed these summary statistics after completion of all six blocks.

Results

Henceforth estimates of effect size for t tests are stated as Cohen's *d*, calculated using averaged standard deviations, and formulae supplied in Howell (1997, p. 217). By convention small, medium and large effect sizes for *d* are .2, .5, and .8 respectively. Estimates of effect sizes for ANOVAs are stated as partial eta-squared (η_p^2). By convention small, medium and large effect sizes for η_p^2 are .01, .06, and .14 respectively.

Discrimination Data

Hit rates (H), the proportion of *Different* responses to Different tone pairs, and false alarm rates (FA), the proportion of *Different* responses to Same tone pairs were calculated for each participant. Each score (H and FA) was derived from a maximum of 36 responses. The data failed to meet the explicit assumptions of signal detection analysis using *d'* (Macmillan & Creelman, 1991). Therefore, each participant's discrimination sensitivity, *A'*, to the 2% difference in tone pitch was calculated using published formulae (Rae, 1976). The means were very similar between participant groups (ASD group: M = .63, SD = .14; TD group: M = .68, SD = .16), with no significant group difference, $t(29)$ = .96, p = .35, d = .35. One sample t tests revealed that both participant groups were discriminating above chance, *A'* = .5 (ASD group: $t(13)$ = 3.59, p < .01, d = .93; TD group: $t(16)$ = 4.76, p < .001, d = 1.13). Following C. Jones et al. (2009), the performance of "exceptionally good discriminators" was defined as an *A'* score that was 1.65 *SD* above the TD group mean. No participant met this criterion. Moderate floor effects likely present in the data would have made it easier to identify any such individuals should they exist. Floor effects reduce the parameters, the standard deviation and the mean, used to identify such individuals.

Accuracy Data

Filler tones were Tones 1, 2, 3, 4, and Tones 7, 8, 9, 10. The ASD group and TD group missed 1.46% and 1.10% of responses to filler tone trials respectively. The total number of correct categorization responses to each filler tone were summed for each participant and expressed as a proportion of unmissed tones (out of a maximum of 96 tones). The response distribution from the TD group was negatively skewed so the data from both participant groups were inverted (subtracted from 1) and transformed with the square root function. Performance on this measure was slightly lower for the ASD group (M = .81, SD = .20) than for the

TD group (*M* = .86, *SD* =.16; reverse transformed scores supplied for means only). However, an independent-samples t test confirmed the difference was not significant, $t(30) = 1.00$, $p = .32$, $d = .35$.

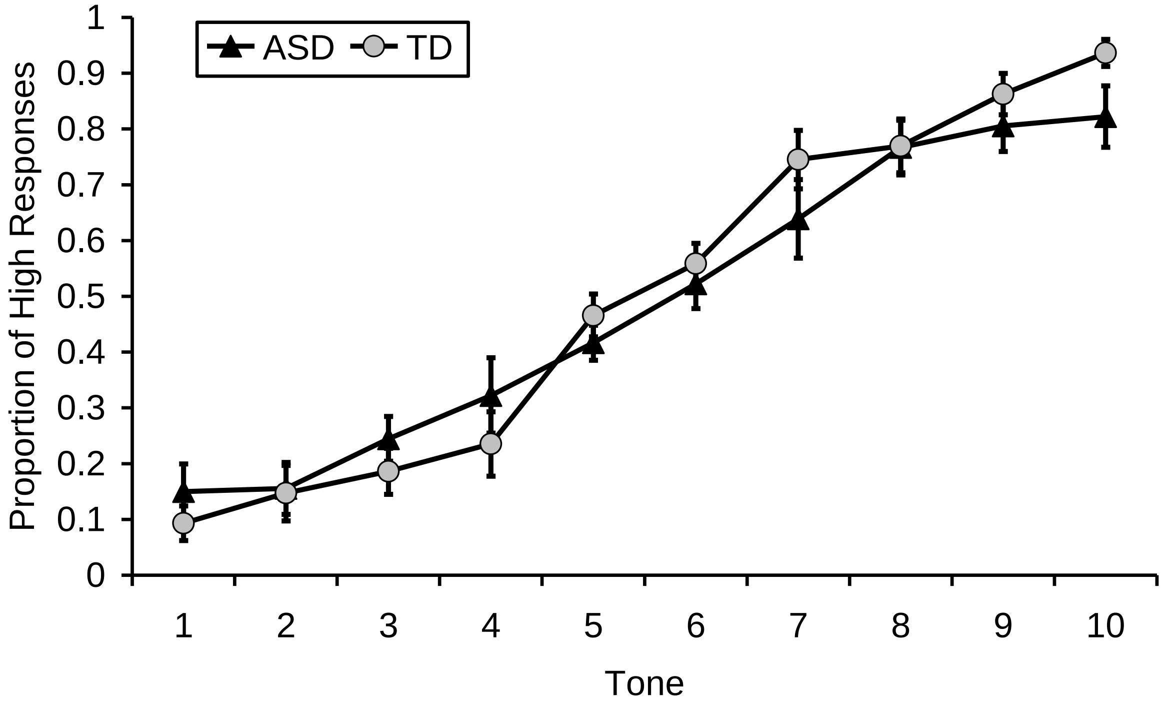

Figure 2. The mean proportion of high responses selected for each tone by each participant group. Error bars represent standard error of the mean.

Figure 2 displays classification plots for each participant group that are very similar. Boundary tones were Tones 5 and 6. The ASD group and TD group missed .83% and .98% of responses to these trials respectively. Figure 3 illustrates the proportion of correct responses (max. = 12) to boundary tones in Same category pairs versus Different category pairs for the ASD group and TD group.

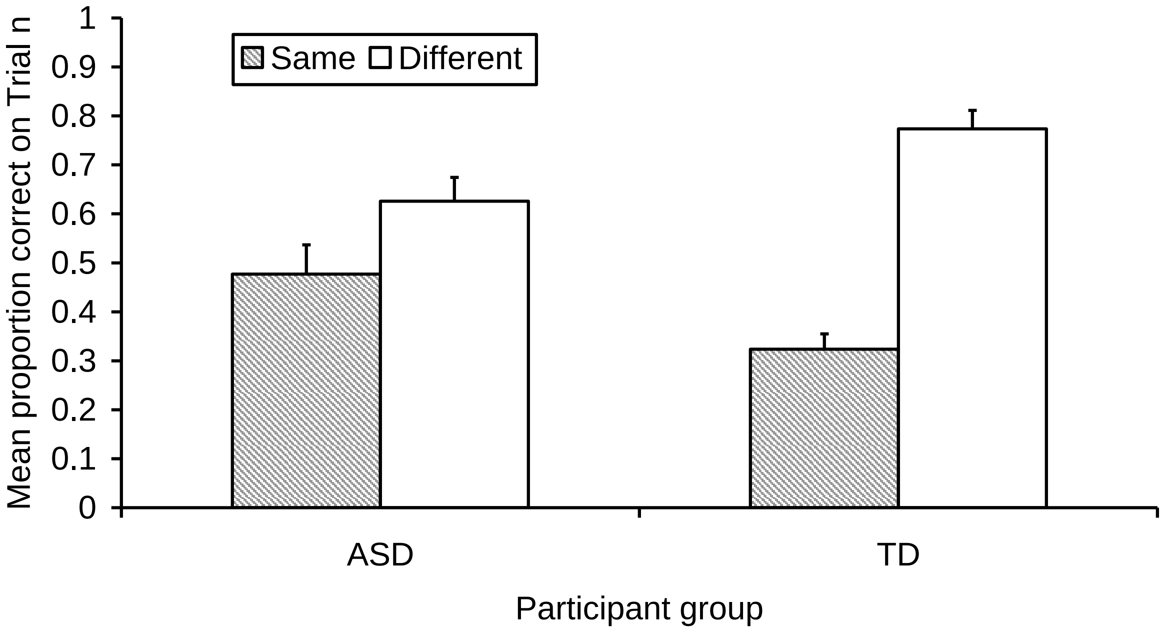

Figure 3. The proportion of correct responses to tones from the Same category (Tones 1 then 5, Tones 10 then 6) and tones from the Different category (Tones 1 then 6, Tones 10 then 5). Error bars represent the standard error of the mean.

Figure 3 illustrates a clear difference in performance between participant groups. The TD group showed a category contrast effect displaying greater accuracy for tones from the Different category than tones from the Same category. The ASD group displayed a greatly reduced effect. The data were analyzed using a mixed repeated measures 2 (group: ASD vs. TD) by 2 (category: Same vs. Different) ANOVA. The group by category interaction was significant, $F(1,30) = 8.37$, $p < .01$, $\eta_p^2 = .22$, as was the main effect of category, $F(1,30) = 33.43$, $p < .001$, $\eta_p^2 = .53$. The main effect of group was not significant, $F(1,30) < 1$, $\eta_p^2 < .001$. Using Bonferroni correction, $\alpha = .025$, the difference in means between categories, Same versus Different, was highly significant for the TD group, $t(16) = 9.15$, $p < .001$, $d = 2.23$, but not for the ASD group, $t(14) = 1.57$, $p = .14$, $d = .41$.[1]

[1] Reaction time data were analyzed in exploratory data analyses but given the small number of data points, no significant effects were observed.

Typical individuals tend to show some perseverative responding (i.e. repeating correct responses from Trial n-1 to Trial n) on unidimensional classification tasks (M. Jones et al., 2006). In addition, some children with autism show marked perseveration on tests of mental flexibility (Hill, 2004). If the participants with autism in this study displayed a greater tendency towards such perseveration this would result in greater accuracy to Same category pairs than to Different category pairs, producing a reverse contrast effect. To explore this possibility, the number of perseverative responses to targets, Tone 4 and Tone 7, were summed and expressed as a proportion of the total unmissed tones (max. = 24, or 23 if a sequence began with a target). Tone 4 and Tone 7 were the most difficult filler tones to categorize and therefore we assumed that responses to these would be the most susceptible to any perseveration effects. The mean proportion of perseverative responses to these tones were very similar between participant groups (ASD group: M = .37, SD = .16; TD group: M = .33, SD = .17) with no significant group difference, $t(30) = .66$, $p = .51$, $d = .23$. Entering the proportion of perseverative responses as a covariate into a 2 (group: ASD vs. TD) by 2 (category: Same vs. Different) ANCOVA made no change to the key finding of a significant group by category interaction, $F(1,29) = 7.63$, $p = .01$, $\eta_p^2 = .21$. This interaction remained strong when the A' score was entered as a covariate instead, $F(1,28) = 12.99$, $p = .001$, $\eta_p^2 = .32$, and when both A' and Perseveration were entered as covariates, $F\ (1,27) = 12.13$, $p = .002$, $\eta_p^2 = .31$. In none of the three ANCOVA reported here were the covariates significant predictors of the contrast effect.

Modeling

To assess the separate influences from long term memory and short term memory, the full set of data for each individual participant were modeled in a logistic regression. With only 120 trials and 68 possible pairings of current tone and preceding tone, the data matrix for each participant was quite sparse. Nonetheless, it was possible to fit a logistic model to each participant's data to predict the response on each trial from three different factors.

The model contained three predictors. The first predictor, Tone, was the tone presented on the current trial, ranging from 1 to 10, reflecting the ability of a participant to compare a tone with a stored criterion in long term memory. The higher the tone the more likely it was predicted that a participant would respond “High”. The second predictor, Distance, was the distance of the current tone from the preceding tone, from +8 to -8. (The extremes, Tones 1 and 10 were not presented to participants in adjacent positions.) Stewart et al. (2002) proposed that instead of using long term memory, participants could successfully categorize the tones to a high level of accuracy by using the distance from the previous tone as a cue for whether to switch the response or not. If the current tone is a lot higher than the previous one, then this provides a cue that the current response should be High (and the same for Low). The two predictors were correlated below 0.7, allowing their independent effects to be estimated. The final predictor, Feedback, was a binary variable reflecting a second strategy proposed by Stewart and colleagues. The predictor was coded as -1 in the case that feedback indicated that the tone on the preceding trial was Low, and the current stimulus to be categorized was either the same as or lower than the preceding one. In such cases, the direction of change of the tone indicates that the response should always be Low. The same variable was coded as +1 for the equivalent High case, where feedback indicated that the preceding tone was High, and the current tone is higher in pitch. On such trials, a “High” response is predicted. On all other trials, the variable was coded as 0. We therefore predicted that if participants are able to use this anchoring strategy, then there should be a positive association between this variable and responding “High” or “Low”. In effect, a code of -1 should predict a “Low” response, and a code of +1 a “High” response, while a code of 0 makes no prediction.

Table 2 shows means and standard deviations for the parameters of the models for each group, together with general goodness of fit measures, the mean R^2 and the mean percentage of responses correctly predicted by the model, across participants. Also shown are the results of a one-sample t-test, testing whether the R^2 and *b* coefficients are greater than zero, and whether the proportion of trials correctly predicted is greater than 50%.

Table 2

Descriptive and test statistics for parameters, Tone, Distance, and Feedback, for each participant group model

	ASD Group				
	R^2	% Correct	*b* Tone	*b* Distance	*b* Feedback
M	0.451	76.907	0.591	0.063	0.293
SD	0.312	13.493	0.531	0.202	0.863
One sample *t*	5.595	7.723	4.309	1.206	1.314
p value	0.001	0.001	0.001	0.248	0.210
95% CI	.28 to .62	69% to 84%	.30 to .89	-0.05 to 0.17	-0.18 to 0.77
d	1.45	1.99	1.11	0.31	0.34
	Typically Developing Group				
	R^2	% Correct	*b* Tone	*b* Distance	*b* Feedback
M	0.559	81.924	0.533	0.198	0.353
SD	0.260	11.000	0.383	0.115	0.784
One sample *t*	8.340	11.240	5.381	6.702	1.745
p value	0.001	0.001	0.001	0.001	0.103

95% CI	.43 to .69	76% to 88%	.33 to .73	.14 to .26	-.05 to .76
d	2.15	2.90	1.39	1.73	0.45

The table clearly supports the analysis based on the boundary tones. Whereas both groups relied quite heavily on the pitch of the tone for their categorization responses (*b* of .591 for the ASD group, and .533 for the TD group), only the TD group showed evidence of using a Distance strategy, based on the degree to which the current stimulus had shifted from the previous one. For the TD group this parameter was significantly greater than zero (*b* = .198, $p < .001$ on one-sample *t*), while for the ASD group it was not significantly different from zero (*b* = .063, $p = .248$). The third predictor, which combined feedback from the previous trial together with the direction of change of the tone was not significant for either group.

Two-sample t-tests applied to the data confirmed that there was no difference in the degree of fit of the models. For R^2, $t(30) = 1.07$, $p = .29$, $d = .38$, and for % Correct, $t(30) = 1.16$, $p = .26$, $d = .4$. Neither was there any difference for the size of the Tone parameter, $t(30) = 0.36$, $p = .72$, $d = .13$. However the Distance parameter was significantly higher for the TD than for the ASD group, in line with the interaction involving boundary tones reported above, $t(30) = 2.37$, $p = .02$, $d = .86$.

Discussion

Children with TD demonstrated a strong category contrast effect demonstrating the influence of sequence effects on their categorization of simple perceptual stimuli. Despite categorization performance that was indistinguishable from that of the TD group, children with autism demonstrated no statistically significant category contrast effect. This finding cannot be attributed to enhanced discrimination or perseveration because it remained after controlling statistically for these variables.

We also found no evidence supporting the idea that reduced category contrast effects could arise from an imbalance in strategy use; weighting resources from long term memory at the expense of those from short term memory. The models we presented, demonstrated that children with autism used long term memory to the same extent as children with TD. Thus, the modeling work used all the data to corroborate the key finding that children with autism are less susceptible to sequence effects. The strategy involving feedback may be too advanced for the children we tested because the related model parameter was not a significant predictor for either participant group.

The MAC model and categorization in autism

The current study demonstrates that the category contrast effect is a highly robust phenomenon. Hitherto, only adult typical populations, mainly university students, have demonstrated the effect. Here we report the effect replicated with a large effect size in a group of schoolchildren with TD.

The MAC model does not account readily for the categorization performance shown by children with autism. According to this model, the sequence effects that drive the category contrast effect, provide the sole basis of categorization performance. Although the contrast effect means for the ASD group were in the right direction, the effect size for the TD group was more than five times that of the ASD group, with no associated superiority in categorization accuracy. Moreover, the model parameter associated with sequence effects, Distance (representing influence from the perceived distance between Tone n and Tone n-1) was a highly significant predictor of categorization performance for the TD group but close to zero and a non-significant predictor for the ASD group. The implication therefore is that participants with autism did in fact use absolute magnitude information from long term memory to categorize given that they did not rely upon sequence effects. Such absolute magnitude information is likely to be *fuzzy* or error prone. If it were not, the binary categorization task would be trivially easy, as Stewart and Brown (2004) argue. Use of even a single accurate representation from the category boundary would result in categorization performance virtually at ceiling. Neither participant group exhibited such performance.

In addition, the autism data illustrate that the mere fact of association, between an influence from sequence effects and successful binary categorization performance, within the TD group does not necessarily imply a causal relationship; successful categorization is clearly achievable with a negligible influence from sequence effects. An unresolved issue therefore is which conditions are necessary before sequence effects go beyond mere epiphenomena to playing an essential role in accurate categorization. It seems likely that sequence effects would assume greatest importance when relevant absolute magnitude information is too fuzzy to be useful, such as at category boundaries or before long term memories are formed, as in the initial stages of categorization. Interestingly, a number of studies using multi-attribute stimuli report initial but temporary weaknesses in autism categorization performance (Bott et al., 2006; Soulières et al., 2011; Vladusich et al., 2010). These findings are expected if sequence effects are less available to participants with autism and if such effects are more important for successful performance during early learning trials.

Discrimination

Enhanced pitch discrimination cannot explain the reduced category contrast effect shown by the ASD group. Contrary to existing research (Bonnel et al., 2003), we found no evidence of enhanced pitch discrimination amongst autism participants at group level. This finding may have arisen from a key methodological difference between the studies. Bonnel et al.'s study entailed repeated presentation of, and discrimination between, only two different stimuli at a time. Enhanced performance reported for the ASD group may reflect a greater capacity for focused attention during this highly repetitive task. A similar explanation has been offered for superior ASD group performance on other highly repetitive perceptual tasks (e.g. Bach & Dakin, 2009). In contrast, the present study presented a greater variety of stimili, which may have promoted TD group performance to ASD group levels. In addition, and contrary to other studies (Heaton et al., 2008; C. Jones et al., 2009), we found no sub-group of enhanced discriminators. Heaton et al. and C. Jones et al. tested larger samples of participants with autism than we did (*n* = 32, and *n* = 71, respectively). Therefore, it seems likely that our failure to recruit enhanced discriminators reflects the smaller sample size of our study.

Weak central coherence

The significantly reduced category contrast effect reported here converges with other evidence suggesting domain general difficulties with contextual processing in autism. This includes the evidence discussed earlier, such as difficulty using context to select a correct pronunciation for a homograph (Frith & Snowling, 1983; Happé, 1997) and difficulty locating an object that is conceptually incongruent with a depicted scene (Jolliffe & Baron-Cohen, 2001). Other convergent evidence includes, for example, difficulties with tests of visual disembedding such as the embedded figures task (Jolliffe & Baron-Cohen, 1997; Shah & Frith, 1983) and the block design task (Happé, 1994b; Shah & Frith, 1993).

A prominent alternative to the weak central coherence account, the "enhanced perceptual functioning" model (Mottron, Dawson, Soulieres, Hubert, & Burack, 2006) suggests that individuals with autism consistently show a local processing bias on hierarchical perceptual stimuli, rather than reduced global processing as suggested by the weak central coherence account. Supporting evidence comes from a number of tasks in which local and global processes are not in competition. For example, autism participants generally show superiority in disembedding a target from an irrelevant surround without any decrement in global processing (Mottron, Peretz, & Ménard, 2000; O'Riordan, Plaisted, Driver, & Baron-Cohen, 2001). The enhanced perceptual functioning model attributes this local bias to the hyper-functioning of low-level auditory or visual perceptual processes. In principle, reduced contextual processing on the homograph task could arise from a local processing bias (Happé & Booth, 2008). In a similar vein, the reduced contextual processing reported here may have arisen from a local bias towards the immediate categorization task. Though this local bias remains a possibility, we found no evidence of enhanced perceptual functioning that in theory should manifest with "a superior perceptual trace" (Mottron et al., 2006, p.28). Such a superior ability should convey advantage for the ASD group on the pitch discrimination and binary categorization tasks and we found none.

Categorization in autism

Studies applying traditional models to study categorization in autism, such as prototype models (Molesworth et al., 2005) and exemplar models (Bott et al., 2006) tend to report typical categorization responses. Such models represent long term and stable category attributes. In contrast, the present findings show a major atypicality; a greatly reduced influence from sequence effects on categorization. This pattern of findings, representing difficulty manipulating dynamic representations from short term memory but intact recall from long term memory, is reflected in other areas of cognitive functioning in autism such as contextual processing on the homograph task described earlier. In addition, studies investigating memory processing in high functioning adults and adolescents with autism reflect this pattern. Poor memory is reported for incidental contextual details surrounding target words such as the gender of the voice presenting the word (Bowler, Gardiner, & Berthollier, 2004). In contrast, information from long term memory stores, such as semantic attributes of target words seems unaffected (Bowler, Gardiner, Grice, & Saavalainen, 2000; Jolliffe & Baron-Cohen, 1999). For example, an ASD group and control group falsely recalled similar proportions of semantic associates (e.g. *mountain*) of target words (e.g. *valley*) studied for a memory test (Bowler et al., 2000).

The MAC model accounts for the ability, found in typical populations, to flexibly update stimulus responses to reflect changing contexts. The ability may convey an adaptive advantage in dynamic environments (M. Jones et al., 2006). A reduction in this ability would give rise to behavioral and cognitive inflexibility. Such inflexibility is characteristic of autism, extending to functioning in both social and non-social domains. For example, individuals with autism may show difficulties shifting problem solving strategy or adopting the viewpoints of others (Geurts, Corbett, & Solomon, 2009). Therefore, the MAC model may uniquely capture the type of categorization process involved in aspects of everyday functioning that are problematic for individuals with autism.

Methodological implications

The current findings have broad methodological implications for any study that presents stimuli in sequence. The implication is that a TD group but not an ASD group will make responses to current trials that reflect the influence of preceding trials. Currently, a variety of stimulus presentation orders tend to be used interchangeably, often without explicit consideration of the effect these choices may have on participant group differences. Such consideration may be necessary to prevent sequence effects acting as a confound, at least for auditory stimuli. Studies investigating the accuracy of estimates of time duration with auditory tones seem to illustrate this point. When stimulus presentation is in fixed ascending order, such that inter-trial differences and hence influences, are minimized, estimates of time duration tend to be indistinguishable between participant groups (Wallace & Happé, 2007). When stimulus presentation order is randomized, so that inter-trial differences are greater and thus potentially more influential, autism participants show less regression towards the median time duration (Martin, Poirier, & Bowler, 2010) and durations at both the shortest and longest extremes exhibit greater participant group differences than median durations (Maister & Plaisted-Grant, 2011). A unified account of time duration estimation in autism could be derived from these data if it is assumed that presentation orders determine the pattern of findings via the differential effect of sequence effects upon autism and control group performance.

Limitations

A limitation of the current study is that we tested a small sample from a clinical group characterized by great heterogeneity (e.g. Geschwind & Levitt, 2007). This means that the findings may be of limited generality within the autism population. The effect sizes reported for all null findings are moderate in places, and some effects might have reached significance with larger sample sizes.

Much of the preceding discussion assumes that the present findings reflect characteristics of domain general processing, however the findings only show effects from auditory stimuli. Analogous studies with a range of stimuli are needed to fully support the domain general conclusions.

Conclusion

The present findings show that children with autism make categorization decisions that are as accurate and efficient as those made by children with TD but with negligible influence from immediately preceding trials. Therefore the MAC model, which assumes that such sequence effects provide the sole basis for categorization, does not readily account for the categorization performance shown by the ASD group. This key finding cannot be attributed to enhanced discrimination within the ASD group because we found no evidence for this. The findings extend the weak central coherence account by demonstrating reduced contextual processing for auditory unidimensional tones. With its focus on dynamic contextual processing, involving short term memory processes, the MAC model may capture an aspect of typical and everyday cognition that is particularly problematic for individuals with autism. Finally, the methodological significance of the current findings is that the influence of preceding trials may act as a potential confound in studies that compare the responses of typical and autism participants to perceptual stimuli presented in sequence.

References

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

Audacity. (Version 1.2.6). [Free computer software] retrieved from http://audacity.sourceforge.net.

Bach, M., & Dakin, S. C. (2009). Regarding "Eagle-Eyed Visual Acuity: An Experimental Investigation of Enhanced Perception in Autism". *Biological Psychology, 66*, 19-20.

Baird, J. C., Green, D. M., & Luce, R. D. (1980). Variability and sequential effects in cross-modality matching of area and loudness. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 277-289.

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.

Bott, L., Brock, J., Brockdorff, N., Boucher, J., & Lamberts, K. (2006). Perceptual similarity in autism. *The Quarterly Journal of Experimental Psychology, 59*, 1-18.

Bowler, D., Gardiner, J. M., & Berthollier, N. (2004). Source memory in adolescents and adults with Asperger's Syndrome. *Journal of Autism and Developmental Disorders, 34*, 533-542.

Bowler, D., 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.

Craig, S., Lewandowsky, S., & Little, D. R. (2011). Error discounting in probabilistic category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 673-687.

E-Prime (Version 2) [Computer software]. Pittsburgh, PA, USA: Psychology Software Tools, Inc. Available from http:// www.pstnet.com.

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

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

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.

Gastgeb, H. Z., Strauss, M. S., & Minshew, N. J. (2006). Do Individuals With Autism Process Categories Differently? The Effect of Typicality and Development. *Child Development, 77*, 1717-1729.

Geschwind, D. H., & Levitt, P. (2007). Autism spectrum disorders: developmental disconnection syndromes. *Current Opinion in Neurobiology, 17*, 103-111.

Geurts, H. M., Corbett, B., & Solomon, M. (2009). The paradox of cognitive flexbility in autism. *Trends in Cognitive Sciences, 13*, 74-82.

Happé, F. (1994a). *Autism: An introduction to psychological theory*. London: UCL Press.

Happé, F. (1994b). Wechsler IQ profile and theory of mind in autism: A research note. *Journal of Child Psychology and Psychiatry, 35*, 1461-1471.

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

Happé, F., & Booth, R. (2008). The power of the positive: Revisiting weak coherence in autism spectrum disorders. *The Quarterly Journal of Experimental Psychology, 61*, 50-63.

Heaton, P., Williams, K., Cummins, O., & Happé, F. (2008). Autism and pitch processing splinter skills. *Autism, 12*, 203-219.

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

Hill, E. (2004). Evaluating the theory of executive dysfunction in autism. *Developmental Review, 24*, 189-233.

Howell, D. C. (1997). *Statistical methods for psychology* (4th ed.). Belmont, CA: Duxbury Press.

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.

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.

Jolliffe, T., & Baron-Cohen, S. (2001). A test of central coherence theory: Can adults with high-functioning autism or Asperger syndrome integrate objects in context. *Visual Cognition, 8*, 67-101.

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

Jones, M., Love, B. C., & Maddox, W. T. (2006). Recency effects as a window to generalization: Separating decisional and perceptual sequential effects in category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 316-332.

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.

Krumhansl, C. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin, 126*, 159-179.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., . . . Rutter, M. (2000). The Autism Diagnostic Observation Schedule--Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*, 205-223.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection Theory: A User's Guide*. New York: Cambridge University Press.

Maister, L., & Plaisted-Grant, K. (2011). Time perception and its relationship to memory in autism spectrum conditions. *Developmental Science, 14*, 1311-1322.

Martin, J. S., Poirier, M., & Bowler, D. M. (2010). Brief report: Impaired temporal reproduction performance in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 40*, 640-646.

Mini Sound Level Meter ST 805 [Apparatus]. (2010). Leeds, UK: Premier Farnell Ltd.

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

Mottron, L., Peretz, I., & Ménard, E. (2000). Local and global processing of music in high-functioning persons with autism. *Journal of Child Psychology and Psychiatry, 41*, 1057-1068.

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

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

Nosofsky, R. M., & Little, D. R. (2010). Classification response times in probabilistic rule-based category structures: Contrasting exemplar-retrieval and decision-boundary models. *Memory & Cognition, 38*, 916-927.

O'Riordan, M., 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.

Rae, G. (1976). Table of A'. *Perceptual and Motor Skills, 42*, 98.

Rimland, B. (1964). *Infantile autism*. New York: Appleton-Century-Crofts.

Scheerer, M., Rothmann, E., & Goldstein, K. (1945). A case of "idiot savant": an experimental study of personality organization. *Psychological Monographs, 58*, 63.

Shah, A., & Frith, U. (1983). An islet of ability in autistic children: A research note. *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. (1982). Geometrical approximations to the structure of musical pitch. *Psychological Review, 89*, 305-333.

Soulières, I., Mottron, L., Giguère, G., & Larochelle, S. (2011). Category induction in autism: Slower, perhaps different, but certainly possible. *The Quarterly Journal of Experimental Psychology, 64*, 311-327.

Stewart, N., & Brown, G. D. A. (2004). Sequence effects in the categorization of tones varying in frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 416-430.

Stewart, N., Brown, G. D. A., & Chater, N. (2002). Sequence effects in categorization of simple perceptual stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 3-11.

Stewart, N., & Matthews, W. J. (2009). Relative judgment and knowledge of the category structure. *Psychonomic Bulletin and Review, 16*, 594-599.

Vladusich, T., Lafe, F., Kim, D., Tager-Flusberg, H., & Grossberg, S. (2010). Prototypical category learning in high-functioning autism. *Autism Research, 3*, 226-236.

Wallace, G., & Happé, F. (2007). Time perception in autism spectrum disorders. *Research in Autism Spectrum Disorders, 2*, 447-455.

Wechsler, D. (1999). *Wechsler abbreviated scales of intelligence*. London: The Psychological Corporation.