Humans spend a lot of their waking time interpreting the actions of others, and to do this rely heavily on visual information. People watch the movements of other people’s eyes in an attempt to infer what is on their minds, and gauge their body language and facial expressions to deduce their mood and intentions. There is good reason to believe that these visually supported social skills are mediated by neural mechanisms specialized for the perception of biological activity (Allison, Puce, & McCarthy, 2000; Grossman et al., 2000; Wheaton, Pipingas, Silberstein, & Puce, 2001). Unfortunately, not all humans are adept at these skills. Individuals with autism, in particular, exhibit chronic deficits in the ability to relate with other people (Kanner, 1943), deficits that may be traceable to difficulties understanding the attitudes and intentions of other people (Baron-Cohen, 1991). To what extent are the deficits in understanding exhibited by individuals with autism traceable to perceptual deficits in their ability to perceive what other people are doing?

Moore, Hobson, and Lee (1997) attempted to answer this question by testing individuals with autism (average age was 14 years) and age-matched control participants on a set of tasks involving recognition of human activity portrayed by point-light animation sequences (Johansson, 1973). With these animations, a dozen or so small “lights” attached to the joints of the body are displayed in brief video sequences, with the motion of the lights tracing the movements of the body. impoverished as they are, these animations readily evoke accurate perception of biological activity, even to the extent that people can judge the activity, sex, and identity of point-light actors (Cutting & Kozlowski, 1977). Moore et al. found that with repeated exposures to these kinds of animations, children with autism could accurately guess whether the sequence portrayed was a human actor or an inanimate object such as an ironing board being opened. In this study, exposure duration was systematically manipulated from brief (40 ms) to long (5 s), and at intermediate durations nonautistic samples consistently performed better than children with autism. These differences failed to reach statistical significance, however, leading Moore et al. to conclude that children and adolescents with autism were not significantly impaired in recognizing that a person was represented in briefly exposed point-light displays. In two other experiments in the study, children with autism performed more poorly than did nonautistic samples, again with the differences being nonsignificant statistically for some but not all comparisons.

Although we have no reason to doubt these results, several considerations led us to believe that the question of biological motion perception in autism remained unresolved. First, one form of motion processing is impaired in autistic children. In particular, two research groups (Milne et al., 2002; Spencer et al., 2000) have reported that autistic children experience relative difficulty perceiving global, coherent motion in random-dot cinematograms: Compared with age-matched normal children, autistic children needed roughly 10% more motion signal before being able to discern a global direction of motion within a field of otherwise random motion. This finding was interpreted as indicative of abnormal functioning within the dorsal stream pathway commonly believed to be involved in motion processing. Now, a deficit in detection of rigid, coherent motion does not necessarily mean that perception of biological motion will likewise suffer. After all, performance levels on these two distinct kinds of tasks are dissociable in normal observers (Grossman & Blake, 1999), in brain-damaged individuals (e.g., Vaina, Lemay, Bienfang, Choi, & Nakayama, 1990), and in people with Williams syndrome (Atkinson et al., 1997; Jordan, Reiss, Hoffman, & Landau, 2002). Still, the fact that autistic individuals are impaired in their ability to detect coherent motion led us to wonder whether they might also have problems perceiving more ecologically relevant events portrayed by motion.

A second reason that motivated us to reexamine the conclusion of Moore et al. (1997) stems from methodological considerations: The task used by Moore et al. may not have been sufficiently sensitive to reveal genuine differences between autistic and nonautistic samples of children. In particular, the behavioral measure consisted of verbal reports in which the child guessed whether each animation portrayed an object or a person (their Experiment 1) or described “what you think the person is doing” (their Experiment 3). These measures are susceptible to response bias (e.g., reports based on expectation, not just sensory data), particularly if the exposure duration is lengthened from trial to trial as the child is being queried. Moreover, the children were not given other, nonbiological motion tasks to assess their general level of attention and motivation.

In view of the potential importance of the conclusions from the Moore et al. (1997) study, we decided to investigate how well children with autism could distinguish biological from nonbiological motion using procedures that are amenable to signal detection analysis, which yields a measure of perceptual sensitivity independent of bias. Furthermore, we tested children on a difficult perceptual grouping task that, although having nothing to do with motion perception, putatively relies on visual mechanisms early in visual processing. We felt this...
task would provide a rigorous assessment of the motivational and attentional capacities of our autistic sample, as well as an index of how well they performed on a task requiring visual grouping of spatially distributed but nonmoving stimulus elements. In addition to measuring group differences between the autistic and nonautistic samples on the two tasks, we tested the relation between task performance and severity of autistic symptoms. Finally, we tested a somewhat younger group of children than did Moore et al., thinking that perhaps perceptual differences might be more pronounced earlier in development, before compensatory strategies have been acquired.

GENERAL METHOD

**Stimuli**

Visual displays were generated on the video monitor of an iMac computer under the control of MatLab© and the Psychophysics Toolbox (Brainard, 1997). All images were rendered in 8-bit gray scale, at a monitor refresh rate of 95 Hz. Viewing distance was 16 in., and the computer provided the only source of illumination in an otherwise darkened room.

**Biological motion task**

Our techniques for generating biological motion sequences have been described elsewhere (Grossman & Blake, 1999). In brief, we video-recorded an adult engaged in a variety of familiar activities, including running, kicking, climbing, throwing, and jumping. We then transcribed those recordings to the computer, placed markers on the joints in each frame of the movie sequence, and then converted those frames to matrices that could be animated and manipulated in MatLab©. Shown in Figure 1 are two frames (not successive in the animation) from a normal biological sequence (in this case, the actor is throwing an object) and two frames (nonsuccessive) from a phase-scrambled sequence created from the “throwing” animation. The phase-scrambled animations consisted of the same individual dots undergoing the same local motions as in the normal animations they were derived from, only with their temporal phases scrambled; this scrambling perturbs the hierarchical, pendular motions characteristic of biological motion, and the resulting animations look distinctly different from their biological counterparts, at least to normal observers.

For formal testing, we created a series of 50 animations, 25 depicting normal biological motion and 25 presenting phase-scrambled sequences; the order of these two types of animations was random within the series. For all animations, the dots appeared black against a gray background (7 cd/m²), and each dot subtended approximately 12 arc min. The duration of each animation exemplar was 1 s. The interframe interval was dictated by the biological stimuli, which required three video retraces to produce the perception of smooth “biological” motion. The average speed within a sequence was about 4 deg/s. The child triggered presentation of each 1-s display, and during the 50-trial sequence the experimenter periodically reminded the child to look in the center of the video screen before starting a trial.

**Global-form task**

Examples of the displays used in this four-alternative, forced-choice task are shown in Figure 2. The entire screen of the video monitor was filled with short lines whose orientations were randomly determined. Each line was approximately 30 min long × 2 min wide, and the lines appeared black against a light gray background. The entire display was divided into four equal-sized quadrants whose boundaries were clearly marked by thick black lines. Within a randomly selected region of one of the quadrants, a small group of eight lines formed a quasi-circular target, and over trials this target could appear in any of the quadrants with equal probability. The clarity of this target was varied by introducing jitter in the orientation of each of the elements forming the circle; jitter was defined as a range of angles within which individual contours defining the target could deviate from the canonical value specified by their position on the circle. Larger degrees of jitter (expressed in angular degrees) produce greater perturbations in the clarity of a target and, hence, would impair participants’ ability to identify in which quadrant the target was located (compare the ease of finding the target in the top and bottom panels of Fig. 2). Our displays

![Fig. 1. Examples of stimuli used in the biological motion task. Two nonsuccessive frames from a point-light animation sequence depicting normal biological activity are shown on the left. On the right are the corresponding frames from an animation containing the same dots undergoing the same local motions, only with their spatiotemporal coherence scrambled to produce meaningless, incoherent motion.](image-url)
were a version of the widely used “pathfinder” display that was devised by Field, Hayes, and Hess (1993) and has been used in the study of visual grouping in normal adults (Kovács & Julesz, 1993), in developing children (Kovács, Kozima, Fehrer, & Benedek, 1999), and in children with Williams syndrome (Atkinson et al., 1997).

**Procedure**

**Biological motion task**

Each child was introduced to the biological motion task by being told he or she would see some short movies of dots that would be either mov-
Autism and Perception of Biological Motion

ing like a person or not moving like a person. Once familiar with the idea of biological motion, each child was tested individually on a 50-trial sequence of 1-s displays, the presentation of which was initiated by the child. Following each presentation, the child verbally reported whether the sequence was “a person” or “not a person.” If a child repeatedly responded with the same answer, the directions were repeated.

Global-form task

Each child was introduced to the global-form task by being asked to find the hidden shape that looked like a “police badge” or a “funny-shaped stop sign.” Sample stimuli were presented before testing began. The testing utilized a four-alternative, forced-choice staircase procedure to find the level of jitter at which the child could correctly identify the quadrant containing the figure on 71% of trials. Testing was preceded by a series of very easy trials in which the circle was perfectly formed (i.e., jitter was zero) and, therefore, easily located. A child was not moved to the test phase until he or she was able to achieve 100% correct performance on a series of 10 of these practice trials.

The actual test was a series of trials on which the figure was made more difficult to detect following each correct answer and less difficult to detect following each incorrect answer. Thus, the child’s performance dictated the trial-by-trial sequence of presentations. Following each response, the correct answer was revealed to the child, providing error feedback. The staircase was terminated after 15 turnarounds, defined as reversals in the direction of the staircase; the mean and standard deviation of the jitter associated with the last 8 turnarounds provided the estimate of threshold performance. Typically staircases lasted about 40 to 50 trials, and the child was given opportunities to rest at any time desired during the staircase. The experimenter remained with the child during testing to ensure that the child’s attention remained focused and that the child did not get discouraged upon making errors.

Participants

Twenty-five children participated in the experiment. Sixteen 8- to 10-year-old children were recruited from a longitudinal study of early diagnosis of autism, and nine 5- to 10-year-old typically developing children were recruited from the community. Four children from the autistic sample were unable to complete the biological motion task, and 3 were unable to complete the global-form task. The children who could not complete the tasks had expressive-language standard scores of 45 or less on the Expressive Vocabulary Test (Williams, 1997). Data for the global-form task from an additional child in the autistic sample were lost because of equipment failure. The results therefore reflect the performance of 12 children from the autistic sample and 9 children from the typically developing sample.

Children in the autistic sample were given a full diagnostic evaluation. They were assessed with a standardized measure of cognitive development (the Kaufman Assessment Battery for Children, or K-ABC; Kaufman & Kaufman, 1983), as well as formal diagnostic measures, including the Autism Diagnostic Observation Schedule-Generic (ADOS-G; Lord et al., 2000) and the Childhood Autism Rating Scale (CARS; Schopler, Reichler, & Renner, 1988). All children were diagnosed by a licensed psychologist (W.L.S.) as meeting diagnostic criteria for Autistic Disorder (American Psychiatric Association, 2000). The chronological age of the typically developing sample (M = 101.0 months, SD = 22.7) was equivalent to the mental age of the autistic sample (M = 94.9 months, SD = 21.3) as measured by the K-ABC, t(20) = 0.64, p = .53. The experimental protocol was approved by the Institutional Research Board of Vanderbilt University, and both parental informed consent and the child’s assent were obtained before participation.

RESULTS

Biological Motion Task

For each child, hits (responding “person” to a biological sequence) and false alarms (responding “person” to a scrambled sequence) were tabulated and used to compute d’, an unbiased measure of sensitivity. The average d’ values for typically developing children and for children with autism are shown in the left-hand panel in Figure 3. The difference between groups was statistically significant, t(19) = 2.68, p = .015.

For the autistic sample, we computed the correlation between severity of autism, as indexed by both the ADOS-G and CARS total scores, and d’ score on the biological motion test. The correlations for both measures were significant, rADOS-G = −.663, p = .019, and rCARS = −.664, p = .018, indicating a significant relationship between severity of autistic symptoms and poor performance on the biological motion test. The d’ scores on the biological motion task were also significantly correlated with mental age in children with autism (r = .75, p = .003), but not with chronological age in typically developing children (r = .24, p = .53).

Given this pattern of results, we cannot rule out the possibility that lower mental age contributes to impairment in performance on the biological motion task, but if it does, this contribution must be specific to children with autism.

Global-Form Task

For each child, we derived a jitter threshold, that is, the range of angular deviations among target contours for which correct identification performance was 71%. The averages of those threshold values are plotted in the right-hand panel in Figure 3. The difference in thresholds for the autistic versus typically developing children was not statistically significant, t(19) = 0.401, p = .69. In contrast to the results for the biological motion task, the performance of children with autism on the global-form task did not correlate with the severity of autistic symptoms, as indexed by either the CARS (r = −.179, p = .58) or the ADOS-G (r = .329, p = .30). Moreover, mental age did not correlate significantly with the visual grouping scores for the children with autism, r = .18, p = .57, nor did chronological age correlate significantly with grouping scores for typically developing children, r = .43, p = .24.

DISCUSSION

Children with autism performed normally on a visual identification task involving the detection of an inanimate object (a “circle”) within a highly cluttered background, even when the contours of that object deviated by 20° to 30° from their canonical orientations. It is important to note that this task is quite challenging: Only a tiny fraction of the line elements within the array formed the figure, and those line elements were indistinguishable from the hundreds of other contours, identical in size and color, forming the background. The only cue specifying the target contours was their spatial arrangement, a cue termed good continuation within the Gestalt tradition. Note, too, that the normal performance of these autistic children is not just another manifestation of their putatively good ability to process parts of stimuli while ignoring potentially distracting contextual features (Happé, 1996; Plaisted,
O’Riordan, & Baron-Cohen, 1998). To perform successfully on our pathfinder task, children had to perceive global regularity in orientation among a subset of neighboring stimulus elements within a field filled with potential figure elements; there was no context to ignore in these displays, as all elements were potentially part of the figure. So, the good performance on this task by the children with autism confirms that they were able to understand the experimenter’s instructions and to maintain an adequate level of attention throughout the session.

As an aside, it is widely believed that the integration of contour information responsible for perception of targets in these pathfinder displays is attributable to neural interconnections among orientation-selective neurons at early stages of visual processing (Field et al., 1993). In fact, there is physiological and anatomical evidence for such connections within visual area V1 (e.g., Gilbert, 1993). Given this interpretation, our results could be interpreted to mean that compromised neurophysiological function in autism spares the primary visual cortex, at least in terms of its integrative circuitry.

Visual perception of biological motion, in contrast, was markedly impaired in our sample of children with autism. We are disinclined to attribute this impairment to poor motivation or wandering attention, for these very same children performed perfectly normally on the form-identification task, which is arguably more difficult and less engaging. One might argue that the differences in performance on the form task and the biological motion task are attributable to differences in developmental trajectories for the two tasks, with optimal performance on the form task being achieved at an earlier age than competence on the biological motion task. Such a difference, the argument continues, could lead to the pattern of results we found because our autistic children were still developing the perceptual expertise needed to perform the biological motion task. However, the data on developmental trends point to exactly the opposite developmental patterns: Children acquire adultlike levels of performance on biological motion tasks at a much earlier age than they do on the form task used here (Fox & McDaniel, 1982; Kovács et al., 1999; Pavlova, Krägeloh-Mann, Sokolov, & Birbaumer, 2001). So, having rejected explanations based on differential motivation between subject groups or on differential developmental trajectories favoring form perception, we are led to conclude that the neural mechanisms responsible for integrating local motion signals into global, coherent biological activity are compromised in autism. Before discussing what those mechanisms might actually be, it is useful to consider our findings in the context of other work on visual behavior in individuals with autism.

Individuals with autism typically exhibit an impaired ability to recognize emotional facial expressions (e.g., Bormann-Kischkel, Vilsmeier, & Baude, 1995). Facial expressions, of course, are conveyed by characteristic patterns of movements of the eyes, forehead, and mouth; they constitute, in other words, a special form of biological motion. It is interesting to note that normal individuals can identify emotional expressions in faces portrayed using point-light animations of the sort employed in our study (Bassili, 1978). Perhaps the difficulties experienced by people with autism in recognizing facial expressions arise, in part, from difficulties integrating motion signals associated with the expression of those facial emotions. In a similar vein, children with autism often fail to shift their gaze when a person they are watching executes head and eye movements indicative of a shift in attention (e.g., Leekam, Baron-Cohen, Perrett, Milders, & Brown, 1997). This deficit, too, could be construed as one involving a breakdown in registration of relevant information about biological motion.

Our biological motion results are also consonant with the idea that individuals with autism experience difficulty integrating component features, or component ideas, into global, coherent patterns (Frith, 1989; Teunisse, Cools, van Spaendonck, Aerts, & Berger, 2001). Termed weak central coherence, this impairment could also underlie the difficulty autistic children have discriminating biological motion from scrambled sequences. In biological motion sequences, individual dots simply undergo translational or elliptical motions. On their own, the individual dot motions carry no information about the human form; such information emerges only from global integration of the local motion signals over space and time. Thus, deficits in global integration—a hallmark of the weak-central-coherence model—could make it difficult to perceive animate activity in these point-light, biological motion animations. At the same time, our autistic children had no trouble integrating small, stationary contours into a global form, a result that is difficult to reconcile with the weak-coherence model. Indeed, results from our form task parallel the findings of other researchers who have found no perceptual performance differences between normal and autistic children (e.g., Ozonoff, Strayer, McMahon, & Filloux, 1994) or even superior performance by autistic children on visuospatial tasks (Plaisted et al., 1998).

![Fig. 3. Mean d' scores on the biological motion task (a) and mean angular threshold of jitter on the global-form task (b) for the autistic and typically developing samples. In both panels, the vertical bars denote ±1 SEM.](image-url)
Autism and Perception of Biological Motion

Much has been written about possible neurological concomitants of autism, and this literature includes some theoretical accounts focusing on disturbances of sensory mechanisms (Ornitz, 1989). Do our findings shed any additional light on this question? Converging lines of evidence point to the involvement of specific regions of the human cerebral cortex in perception of visual information signaling the activities and intentions of other humans (Allison et al., 2000). Located in and near the superior temporal sulcus (STS), these regions contain neurons that are selectively activated by visualization of movements of the face, head, eyes, and body of an individual. In monkeys this area has been studied using single-cell recording techniques (e.g., Perrett et al., 1985), and in humans area STS can be reliably identified using functional magnetic resonance imaging (fMRI) to measure neural activation upon presentation of biological motion displays (e.g., Grossman et al., 2000). Moreover, brain damage in the STS region impairs the ability to recognize biological motion animations but spares other aspects of motion perception (Schenk & Zihl, 1997a, 1997b).

Given these results and conclusions, one naturally wonders whether area STS is functionally compromised in autism, thereby adversely affecting the abilities of individuals with autism to judge the intentions of others. It may be noteworthy that the STS sends signals to and receives signals from the amygdala, a limbic structure thought to attach affective significance to sensory information. Autism, as we pointed out earlier, is also characterized by disturbances in the ability to judge emotional reactions in other people. Of course, we have no idea about the etiology of this possible link between the STS and impaired perception of biological motion. The primary deficit could arise within the limbic system, which, in turn, fails to attach normal emotion significance to STS signals about biological motion and, thereby, adversely affects neural development of the STS. This causal chain of events would be consistent with the recently advanced hypothesis that a general lack of social interest in young autistic children leads to stunted development of cortical mechanisms responsive to the human form (Grelotti, Gauthier, & Schultz, 2002; Schultz et al., 2000). Alternatively, children born with autism might have congenital deficits in neural processing within the STS, independent of this area’s connectivity with other brain regions or of the child’s level of social motivation. Our results do not distinguish between these alternatives.

As mentioned in the introduction, compared with normally developing children, autistic children require a larger percentage of dots moving in a given direction in order to detect coherent motion within a field of other dots moving in random directions (Milne et al., 2002; Spencer et al., 2000). Of course, biological motion tasks are a quite different kind of task, in that the animations comprise very few dots and their “coherence” is defined by the kinematics of human motion, not by rigid, translational motion. Still, the existence of deficits on both kinds of motion tasks, similar though they are, further strengthens the conjecture that the dorsal stream pathway (which includes visual area STS) is impaired in autism.

Whatever the neural bases of the deficits we and other researchers have documented, our results serve as a reminder that impairment in social function, which can lead to withdrawal, may have at least some of its roots in perceptual disorders.

Acknowledgments—We are grateful to the children and parents who participated in this research, to Sarah McGrath and Lynn Kubos for help in the assessment process, and to Elaine Coonrod for comments on an earlier draft of the manuscript. This research was supported in part by grants from the National Institutes of Health (EY07760, NICHD T3207226) and the Hobbs Society of the John F Kennedy Center for Research in Human Development.

REFERENCES


(RECEIVED 12/24/01; REVISION ACCEPTED 6/12/02)