

## Brief Report: Can You See What is Not There? Low-level Auditory–visual Integration in Autism Spectrum Disorder

Maarten J. van der Smagt · Herman van Engeland ·  
Chantal Kemner

Published online: 2 February 2007  
© Springer Science+Business Media, LLC 2007

**Abstract** Patients diagnosed with Autism Spectrum Disorder, show impaired integration of information across different senses. The processing-level from which this impairment originates, however, remains unclear. We investigated low-level integration of auditory and visual stimuli in subjects with Autism Spectrum Disorder. High-functioning adult subjects with Autism Spectrum Disorder as well as age- and IQ-matched adults were tested using a task that evokes illusory visual stimuli, by presenting sounds concurrently with visual flashes. In both groups the number of sounds presented significantly affected the number of flashes perceived, yet there was no difference between groups. This finding implicates that any problems arising from integrating auditory and visual information must stem from higher processing stages in high-functioning adults with Autism Spectrum Disorder.

**Keywords** Autism · High-functioning · Cross-modal integration · Visual perception · Auditory perception · Illusions

---

M. J. van der Smagt (✉)  
Experimental Psychology, Helmholtz Institute & Utrecht  
University, Heidelberglaan 2, Utrecht 3584 CS,  
The Netherlands  
e-mail: M.J.vanderSmagt@uu.nl

H. van Engeland · C. Kemner  
Rudolf Magnus Institute of Neuroscience, Department of  
Child and Adolescent Psychiatry, University Medical  
Center Utrecht, Utrecht, The Netherlands

C. Kemner  
Department of Neurocognition, Faculty of Psychology,  
Maastricht University, Maastricht, The Netherlands

Autism Spectrum Disorder (ASD) is a developmental disorder, of which autism is the most severe form. Recent studies suggest that there are widespread neurodevelopmental abnormalities in ASD that might be related to the integration of information from multiple brain regions (Barnea-Goraly et al., 2004; Bertone, Mottron, Jelenic, & Faubert, 2003; Cherkassky, Kana, Keller, & Just, 2006; Just, Cherkassky, Keller, Kana, & Minshew, 2006; Just, Cherkassky, Keller, & Minshew, 2004). Also, it has been argued that more general perceptual atypicalities in ASD might be related to abnormalities in sensory integration (see Iarocci & McDonald, 2006) for a review).

A specific type of sensory integration is that in which information from different modalities (such as visual and auditory) is combined. Multi-sensory integration of visual and auditory information is particularly relevant for social situations, such as the perception of emotions and language. For instance lip-reading can improve speech understanding, mainly under conditions of poor auditory intelligibility, as in noisy environments (Sumby & Pollack, 1954). Similarly, decreased behavioral response latencies are found for bimodal versus unimodal recognition of emotions (de Gelder & Vroomen, 2000). Therefore, this multi-sensory integration is especially relevant to the problems of language and emotion processing shown by subjects with ASD.

Muller, Kleinhaus, Kemmotsu, Pierce, & Courchesne (2003) demonstrated abnormal fMRI activation patterns in subjects with autism in a task that required the integration of visual and motor information (on which performance was impaired, compared to control subjects). Autistic subjects also performed worse on an emotion recognition task

involving the combined use of visual and auditory information, and showed concurrent abnormal cerebral blood flow patterns (Hall, Szechtman, & Nahmias, 2003). A diminished (facilitatory) effect of visual speech on auditory speech perception was found in children with ASD (functioning in normal IQ-range) compared to healthy controls (de Gelder, Vroomen, & van der Heide, 1991).

There is considerable evidence that sensory integration occurs in specific brain areas that are sensitive to information from different sensory modalities (e.g. Stein & Meredith, 1993). ERP studies indicate that this processing usually occurs later in time, and is therefore associated with higher order processing (see e.g. Klucharev, Mottonen, & Sams, 2003; Lebib et al., 2004). However, there is accumulating evidence that multimodal integration also includes the modulation of activity at cortical brain sites that used to be considered modality specific and are usually related to perceptual aspects of processing (Calvert et al., 1997, 1999). Most studies in subjects with ASD used audio–visual stimuli that implicated higher level (more cognitive) processing, like in the studies on speech and emotion processing (de Gelder et al., 1991; Hall et al., 2003). Therefore, it is unclear to what extent the results of these tasks reflect *perceptual* aspects of (abnormal) multimodal integration.

Recently, multimodal integration has been demonstrated in healthy subjects in a task with much simpler stimuli (Shams, Kamitani, & Shimojo, 2000). In this task, visual flashes are presented, and subjects are requested to count these. Sounds (short transient beeps) are presented concurrently with the visual flashes and evoke additional, illusory flashes; the number of presented beeps influences the number of flashes perceived (Shams et al., 2000). EEG activity measured during the task has shown that the perception of illusory flashes concurs with increased *early* EEG activity above the visual cortex (Shams, Kamitani, Thompson, & Shimojo, 2001) indicating auditory–visual integration at a low, *sensory*, level. Additional evidence comes from a study by Arden, Wolf, and Messiter (2003), who showed that sound alone does not drive primary visual cortex (V1), yet the combination of the auditory and visual stimuli triggers additional activity in V1, which may drive the illusion. Since the duration between the sound activating an already primed visual cortex is in the order of 20–45 ms, this indicates auditory–visual integration on a low (sensory) level. Moreover, the fact that the illusion occurs (even in non-naïve observers) indicates that it reflects a bottom-up process, over which subjects have no voluntary or attentional control.

In the present study this illusion is used to test low-level auditory–visual integration in high-functioning adults with ASD. Abnormal multimodal integration in subjects with ASD *at this level* should result in a decrease in the occurrence or strength of the illusion compared to the normal controls. On the other hand, normal performance of subjects with ASD would indicate that possible problems with auditory–visual integration have to originate from higher (cognitive) processing levels.

## Methods

### Subjects

Fifteen individuals with ASD and fifteen healthy control individuals (13 males, 2 females in each group), matched for age and IQ (see Table 1) participated in the study.

The clinical subjects were recruited via the Department of Child and Adolescent Psychiatry at the University Medical Center in Utrecht, the control subjects from schools for higher education in Utrecht. The study was described to the subjects and written informed consent was obtained according to the Declaration of Helsinki and as approved by the Ethical Committee of the University Medical Center in Utrecht. Diagnoses of either Autistic Disorder or Asperger Syndrome were based on DSM-IV criteria (American-Psychological-Association, 1994). Also, the parents of all autistic subjects were administered the Autism Diagnostic Interview Revised (ADI-R) (Lord, Rutter, & Le Couteur, 1994), and the Autism Diagnostic Observation Scale (ADOS) (Lord et al., 2000) was obtained from the autistic subjects, by certified raters. Eight subjects met the full criteria for autism on both scales, while the remaining seven met the full criteria for autism on either ADI-R or ADOS and fell one point short of meeting criteria on the

**Table 1** Mean age, total IQ, verbal IQ, and performal IQ for both subjects with ASD and control subjects

	Controls		ASD	
	Mean	SD	Mean	SD
Age (year)	20.7	2.6	20.5	3.2
Total IQ	119	11	122	11
Verbal IQ	121	14	123	13
Performal IQ	113	14	116	11

other (thereby fulfilling criteria for ASD; see Table 2).

### Experimental Conditions

The stimuli were generated on an Apple G4 computer using Matlab and the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). Visual stimuli were presented on a LaCie electronblue IV 22 inch monitor. Auditory stimuli were presented through standard external computer speakers, positioned adjacent (left and right) to the monitor.

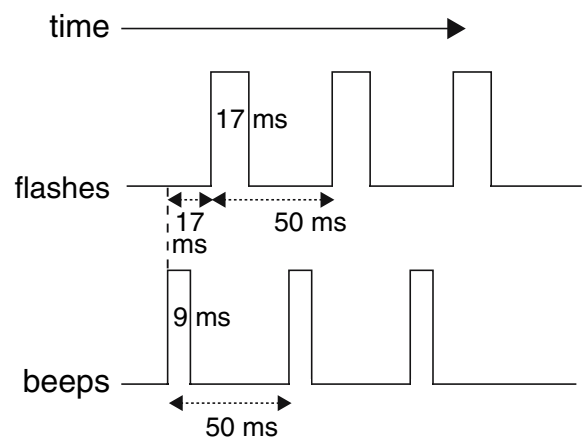
The experiment consisted of 12 conditions: 3 visual conditions (either 1, 2 or 3 visual flashes), combined with 4 auditory conditions (0, 1, 2, or 3 beeps), and thus included consistent (with the same number of beeps and flashes) as well as inconsistent trials. The conditions with 1 visual flash were the crucial conditions, as these are similar to those that produced the illusion in the (Shams et al., 2000) study. The conditions with 2 and 3 flashes serve to control for the possibility that subjects ignore the visual stimuli altogether and only respond to the auditory stimulus. The conditions without auditory stimuli (0 beep conditions) are included to confirm that the subjects are able to distinguish between the three visual conditions used.

The visual stimulus was a white disk ( $46 \text{ cd/m}^2$ ) subtending a visual angle of  $2^\circ$ , displayed on a dark background ( $1 \text{ cd/m}^2$ ),  $6^\circ$  left or right from a central fixation cross. The presentation side was randomly varied from trial to trial, in order to ensure that subjects were not tempted to shift fixation towards the stimulus location. The presentation duration of the disc was 17 ms. If multiple flashes were presented, the stimulus onset asynchrony (SOA) between flashes was 50 ms. The auditory stimulus consisted of one or more beeps (3.5 kHz,  $\sim 75 \text{ dB SPL}$ ) lasting 9 ms and with a SOA of 50 ms. The SOA between the first beep and the onset of the first flash was 17 ms (see Fig. 1).

There were 15 randomly presented trials for each condition. Participants were asked to indicate, by pressing keys 1–3 on a numeric keypad, how many flashes they perceived.

**Table 2** Number of subjects ( $n$ ) meeting criteria for autism or autism spectrum on both ADI-R and ADOS scales (Lord et al., 1994, 2000)

$n$	Diagnostic Scale	
	ADI-R	ADOS
8	Autism	Autism
5	Autism	Autism-spectrum
2	Autism-spectrum	Autism



**Fig. 1** Temporal profile of the stimuli used in the experiments. The 3 flashes—3 beeps condition is depicted. The other conditions are identical but contain fewer flashes or beeps. For instance, the 2 flash conditions contain only the first two flashes

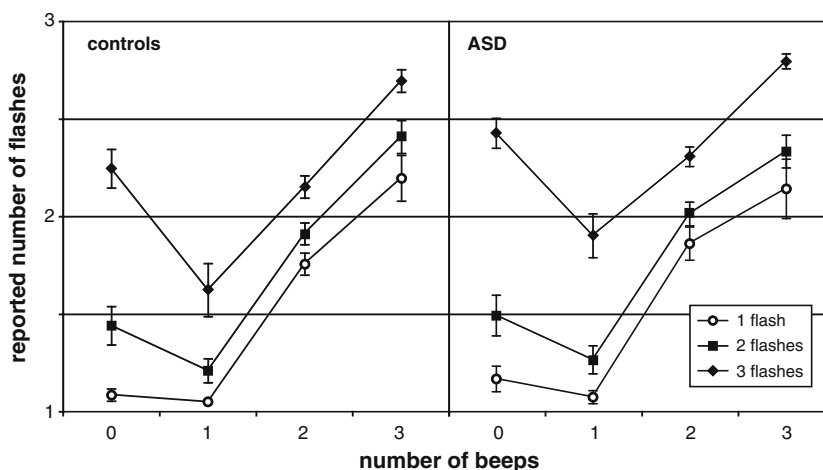
### Results

The results for both groups are presented in Fig. 2, where the number of reported flashes are plotted as a function of the number of presented beeps for the controls (left panel) as well as subjects with ASD (right panel). The parameter is the number of presented flashes (1, open circles, 2, closed squares and 3, closed diamonds).

It is immediately clear from the figure that the subjects did not perform veridical on a purely visual task, and that this holds for both controls and subjects with ASD. It is also apparent from the figure that when multiple flashes were accompanied by a single beep, the reported number of flashes is decreased compared to both the multiple beep conditions and zero beep conditions across groups (bonferroni corrected paired samples  $t$ -tests across groups:  $T > 3.8$ ,  $p < 0.01$  for all comparisons). Both these results are at odds with the results presented by Shams et al. (2000), and will be briefly discussed below.

The main question of this study, however, was to address whether subjects with ASD show normal or abnormal auditory–visual integration, as compared to the control subjects. To analyze these results statistically, a repeated measures analysis of variance was used. Increasing the number of flashes or the number of beeps resulted in an increase in the number of perceived flashes (main effect of flashes:  $F(2,27) = 93.8$ ,  $p < 0.001$ ; main effect of beeps:  $F(3,27) = 77.8$ ,  $p < 0.001$ ). A significant interaction between the number of flashes and number of beeps was also found ( $F(6,23) = 18.7$ ,  $p < 0.001$ ). The number of beeps presented thus significantly affected the

**Fig. 2** Effects of the number of presented beeps on the reported number of flashes for both controls and subjects with ASD (means and standard errors of the mean). Open circles represent the condition when a single visual flash was presented. Closed squares and diamonds represent the conditions with two and three visual flashes respectively



number of flashes perceived. The flashes\*beeps interaction found, however, appears to mainly reflect the relatively larger difference between 3 flashes and 1- and 2-flashes conditions when no beeps were presented compared to the conditions with one or more beeps. The crucial outcome of the experiment is that no interaction whatsoever was found with the factor group (flashes\*group:  $p > 0.16$ , beeps\*group:  $p > 0.18$  and flashes\*beeps\*group:  $p > 0.71$ ). To test as sensitive as possible for any difference between the groups, post-hoc  $t$ -tests (*not* corrected for multiple comparisons) were done for each of the 12 conditions, but none were found.

## Discussion

In research on Autism Spectrum Disorder (ASD) there is increasing focus on the ability to integrate the output of different brain areas (Barnea-Goraly et al., 2004; Bertone et al., 2003). We studied integration of auditory–visual information in high-functioning young adults with ASD, using a task in which an auditory stimulus invokes the perception of an illusory visual stimulus (Shams et al., 2000).

The results of our experiments differ from those of Shams et al. (2000) in two ways. First, performance in the conditions without sound is non-veridical for either group. This can be explained in part by the fact that in the present experiment only the percept of 1, 2 or 3 flashes could be reported. As a consequence, any incorrect response to, for instance, the 3 flashes condition results in the report of a lower number of flashes. Some ‘compression’ of the data is thus expected. In the Shams et al. (2000) study up to 4 visual flashes could be reported, and a similar compression is also apparent from their data on the

4 flashes condition. In addition, the visual stimulus in the present experiment was randomly positioned  $6^\circ$  either left or right from fixation, while in the original study it was always positioned in a single location ( $6^\circ$  below fixation), making the present visual task a more difficult one. This might also explain the second difference between the two studies. A harder visual task will cause the auditory stimulus to have a more profound effect, hence the fact that a single beep in the present experiment decreases the number of reported (multiple) flashes, while this effect was absent in the Shams et al. (2000) study. Despite these small differences, the results from the main experimental (single flash) conditions were remarkably similar to those previously reported.

The expected illusory effect was found in both the control and clinical groups: the number of concurrently presented sounds influenced the number of flashes perceived. These results indicate that the subjects with ASD *did* integrate the auditory and visual information, probably at an early (sensory) level of processing. Our findings are in accordance with the results from a recent study, indicating normal discrimination of temporal synchrony in non-linguistic intermodal stimuli in mentally retarded young children with ASD (Bebko, Weiss, Demark, & Gomez, 2006) and studies indicating that patients with ASD show normal integration of visual and auditory speech stimuli (e.g. Williams, Massaro, Peel, Bosseler, & Suddendorf, 2004). This implicates that, although abnormalities in white matter tracts involved in integration of information between different brain areas have been found (Barnea-Goraly et al., 2004), at least some connections between auditory and visual brain areas appear to function appropriately in ASD.

The illusory effect detected in the present study is thought to be caused by auditory brainstem activity

that passes through the thalamic radiation to the primary visual cortex (Arden et al., 2003). Possible involvement of sub-cortical, especially thalamic, structures in this task is noteworthy because two studies of multimodal integration found evidence of abnormal thalamic activity in subjects with ASD. Abnormal thalamic activation was found in subjects with autism during auditory–visual integration of emotional cues (Hall et al., 2003), and in a study of visuo-motor integration subjects with autism showed abnormal activation patterns which were hypothesized to be related to developmental disturbances in thalamo-cortical afferents (Muller et al., 2003). Moreover, there are indications that the thalamus is smaller in men with high-functioning autism than in normal control men (Tsatsanis et al., 2003). Surprisingly, a recent study indicated more extensive thalamo-cortical functional connectivity in high functioning men with autism, compared to controls, which is in contrast to the hypothesis of general underconnectivity in ASD (Mizuno, Villalobos, Davies, Dahl, & Muller, 2006). The apparently normal performance of subjects with ASD in the present study does not support the idea that abnormalities in thalamic functioning play an important role in the potential problems with auditory–visual integration shown by high-functioning individuals with ASD, although different pathways might be involved. For instance, it has been suggested that autistic adults may use the non-classical auditory pathways (Moller, Kern, & Grannemann, 2005), which are known to regress with age in healthy individuals (Moller & Rollins, 2002). Obviously, the present results do not exclude abnormal multimodal integration at later processing stages, which involve other brain areas, such as the superior temporal sulcus (STS) (e.g. Boddaert et al., 2004).

It should be noted that the precise mechanisms of the auditory–visual illusion are not yet established. In addition, the subjects in the present study are a distinct group of high functioning young adults, and abnormal auditory–visual integration might be present in other (low-functioning or younger) subjects with ASD. However, the present data suggest that, at least in high-functioning adults with ASD, any problems in domains of functioning that rely on both visual and auditory information, such as emotion and language processing, are not likely to be the result of abnormal low-level auditory–visual integration.

**Acknowledgments** The work described was supported by Innovational Research Incentives grants (VENI-scheme, 451-02-094) to Maarten J. van der Smagt, and (VIDI-scheme, 402-01-094) to Chantal Kemner, of the Netherlands Organisation for Scientific Research (NWO).

## References

- American-Psychological-Association. (1994). *Diagnostic and statistical manual of mental disorders*. Washington, DC: APA.
- Arden, G. B., Wolf, J. E., & Messiter, C. (2003). Electrical activity in visual cortex associated with combined auditory and visual stimulation in temporal sequences known to be associated with a visual illusion. *Vision Research*, *43*, 2469–2478.
- Barnea-Goraly, N., Kwon, H., Menon, V., Eliez, S., Lotspeich, L., & Reiss, A. L. (2004). White matter structure in autism: Preliminary evidence from diffusion tensor imaging. *Biological Psychiatry*, *55*, 323–326.
- Bebko, J. M., Weiss, J. A., Demark, J. L., & Gomez, P. (2006). Discrimination of temporal synchrony in intermodal events by children with autism and children with developmental disabilities without autism. *Journal of Child Psychology and Psychiatry*, *47*, 88–98.
- Bertone, A., Mottron, L., Jelenic, P., & Faubert, J. (2003). Motion perception in autism: A “complex” issue. *Journal of Cognitive Neuroscience*, *15*, 218–225.
- Boddaert, N., Chabane, N., Gervais, H., Good, C. D., Bourgeois, M., Plumet, M. H., et al. (2004). Superior temporal sulcus anatomical abnormalities in childhood autism: A voxel-based morphometry MRI study. *Neuroimage*, *23*, 364–369.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.
- Calvert, G. A., Brammer, M. J., Bullmore, E. T., Campbell, R., Iversen, S. D., & David, A. S. (1999). Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport*, *10*, 2619–2623.
- Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire, P. K., et al. (1997). Activation of auditory cortex during silent lipreading. *Science*, *276*, 593–596.
- Cherkassky, V. L., Kana, R. K., Keller, T. A., & Just, M. A. (2006). Functional connectivity in a baseline resting-state network in autism. *Neuroreport*, *17*, 1687–1690.
- de Gelder, B., & Vroomen, J. (2000). Bimodal emotion perception: Integration across separate modalities, cross-modal perceptual grouping or perception of multimodal events? *Cognition & Emotion*, *14*, 321–324.
- de Gelder, B., Vroomen, J., & van der Heide, L. (1991). Face recognition and lip-reading in autism. *European Journal of Cognitive Psychology*, *3*, 69–86.
- Hall, G. B., Szechtman, H., & Nahmias, C. (2003). Enhanced salience and emotion recognition in Autism: A PET study. *American Journal of Psychiatry*, *160*, 1439–1441.
- Iarocci, G., & McDonald, J. (2006). Sensory integration and the perceptual experience of persons with autism. *Journal of Autism and Developmental Disorders*, *36*, 77–90.
- Just, M. A., Cherkassky, V. L., Keller, T. A., Kana, R. K., & Minshew, N. J. (2006). Functional and anatomical cortical underconnectivity in autism: Evidence from an fMRI study of an executive function task and corpus callosum morphometry. *Cerebral Cortex*.
- Just, M. A., Cherkassky, V. L., Keller, T. A., & Minshew, N. J. (2004). Cortical activation and synchronization during sentence comprehension in high-functioning autism: Evidence of underconnectivity. *Brain*, *127*, 1811–1821.
- Klucharev, V., Mottonen, R., & Sams, M. (2003). Electrophysiological indicators of phonetic and non-phonetic multisensory interactions during audiovisual speech perception. *Brain Research*, *18*, 65–75.



- Lebib, R., Papo, D., Douiri, A., de Bode, S., Gillon Dowens, M., & Baudonniere, P. M. (2004). Modulations of 'late' event-related brain potentials in humans by dynamic audiovisual speech stimuli. *Neuroscience Letters*, *372*, 74–79.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H. Jr., Leventhal, B. L., DiLavore, P. C., et al. (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.
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*, 659–685.
- Mizuno, A., Villalobos, M. E., Davies, M. M., Dahl, B. C., & Muller, R. A. (2006). Partially enhanced thalamocortical functional connectivity in autism. *Brain Research*, *1104*, 160–174.
- Moller, A. R., Kern, J. K., & Grannemann, B. (2005). Are the non-classical auditory pathways involved in autism and PDD? *Neurological Research*, *27*, 625–629.
- Moller, A. R., & Rollins, P. R. (2002). The non-classical auditory pathways are involved in hearing in children but not in adults. *Neuroscience Letters*, *319*, 41–44.
- Muller, R. A., Kleinhans, N., Kemmotsu, N., Pierce, K., & Courchesne, E. (2003). Abnormal variability and distribution of functional maps in autism: An fMRI study of visuomotor learning. *American Journal of Psychiatry*, *160*, 1847–1862.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature*, *408*, 788.
- Shams, L., Kamitani, Y., Thompson, S., & Shimojo, S. (2001). Sound alters visual evoked potentials in humans. *Neuroreport*, *12*, 3849–3852.
- Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. Cambridge, Massachusetts: The MIT Press.
- Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, *26*, 212–215.
- Tsatsanis, K. D., Rourke, B. P., Klin, A., Volkmar, F. R., Cicchetti, D., & Schultz, R. T. (2003). Reduced thalamic volume in high-functioning individuals with autism. *Biological Psychiatry*, *53*, 121–129.
- Williams, J. H., Massaro, D. W., Peel, N. J., Bosseler, A., & Suddendorf, T. (2004). Visual-auditory integration during speech imitation in autism. *Research in Developmental Disabilities*, *25*, 559–575.