



Liu, P., Sutherland, M. and Pollick, F. E. (2021) Incongruence effects in cross-modal emotional processing in autistic traits: an fMRI study. *Neuropsychologia*, 161, 107997. (doi: [10.1016/j.neuropsychologia.2021.107997](https://doi.org/10.1016/j.neuropsychologia.2021.107997))

The material cannot be used for any other purpose without further permission of the publisher and is for private use only.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/249955/>

Deposited on 30 August 2021

Enlighten – Research publications by members of the University of
Glasgow

<http://eprints.gla.ac.uk>

Title: Incongruence effects in cross-modal emotional processing in autistic traits: An fMRI study

Author names and affiliations:

Peipei Liu ^{a,b,c}, Margaret Sutherland ^c, Frank E. Pollick ^{b,*}

^a Department of Psychology, Sun Yat-Sun University, Guangzhou, 510006,, UK

^b School of Psychology, University of Glasgow, Glasgow, G12 8QB, UK

^c School of Education, University of Glasgow, Glasgow, G3 6NH, UK

***Correspondence:**

Professor Frank E. Pollick

School of Psychology,

University of Glasgow,

62 Hillhead St.

Glasgow, G12 8QB

E-Mail: Frank.Pollick@glasgow.ac.uk

Abstract

In everyday life, emotional information is often conveyed by both the face and the voice. Consequently, information presented by one source can alter the way in which information from the other source is perceived, leading to emotional incongruence. Here, we used functional magnetic resonance imaging (fMRI) to examine neural correlates of two different types of emotional incongruence in audiovisual processing, namely incongruence of emotion-valence and incongruence of emotion-presence. Participants were in two groups, one group with a low Autism Quotient score (LAQ) and one with a high score (HAQ). Each participant experienced emotional (happy, fearful) or neutral faces or voices while concurrently being exposed to emotional (happy, fearful) or neutral voices or faces. They were instructed to attend to either the visual or auditory track. The incongruence effect of emotion-valence was characterized by activation in a wide range of brain regions in both hemispheres involving the inferior frontal gyrus, cuneus, superior temporal gyrus, and middle frontal gyrus. The incongruence effect of emotion-presence was characterized by activation in a set of temporal and occipital regions in both hemispheres, including the middle occipital gyrus, middle temporal gyrus and inferior temporal gyrus. In addition, the present study identified greater recruitment of the right inferior parietal lobule in perceiving audio-visual emotional expressions in HAQ individuals, as compared to the LAQ individuals. Depending on face or voice-to-be attended, different patterns of emotional incongruence were found between the two groups. Specifically, the HAQ group tend to show more incidental processing to visual information whilst the LAQ group tend to show more incidental processing to auditory information during the crossmodal emotional incongruence decoding. These differences might be attributed to different attentional demands and different processing strategies between the two groups.

Keywords: fMRI, emotional conflict, incongruence, autism quotient, face-voice

1. Introduction

Autism spectrum disorder (ASD) defines an early-onset, heterogeneous neurodevelopmental condition, characterized by the presence of difficulties with social interactions and communication together with restricted or repetitive interests and behaviors (American Psychiatric Association, 2013). As the core feature of ASD, difficulties in social communication are often linked to differences in emotion processing, leading to difficulties in understanding and appreciating social information in real-life situation. Previous studies using unimodal stimuli have demonstrated that individuals with ASD tend to show abnormalities in attending to and extracting salient information from social cues including facial expressions, voice prosody and eye gaze, compared to typical developing population (Dawson et al., 1998; Hobson et al., 1989; Klin, 1991; Rutherford, 2007). At the neurobiological level, these atypical patterns of emotion recognition have been linked to processing differences in brain areas such as fusiform gyrus (FG) (e.g., face processing), superior temporal sulcus (STS) (e.g., voice processing) and medial prefrontal cortex (mPFC) (Bozidakas et al., 2006; Gervais et al., 2004; Hall et al., 2003; Happé et al., 1996; Nieminen-von Wendt et al., 2003). Consequently, individuals with ASD often experience difficulty in understanding social information in situations of daily life.

In natural social interactions, emotions are generally expressed and perceived through multiple domains. Appreciating multi-sensory information is likely to require both attending and extracting relevant information from different sources and integrating this with the context of the situation at hand (Campanella and Belin, 2007). A few studies have looked into the localisation of brain regions involved in audio-visual processing in ASD and tend to report functional neurological differences underlying this process. For example, there is research suggesting that

ASD individuals tend to show hypo activation in emotion and integrative regions in frontal and temporal lobes, such as inferior frontal cortex (IFC) (Hall et al., 2003), mPFC (Wang et al., 2007), fronto-limbic areas (Loveland et al., 2008), superior temporal gyrus (STG) (Loveland et al., 2008; Wang et al., 2007), and FG (Hall et al., 2003). Notably, despite successful behavioral performance on an audiovisual emotion matching task, Doyle-thomas et al. (2013) reported that in contrast to more involvement of frontal and temporal associative cortices in neurotypicals, individuals with ASD show more activation in the parietofrontal network. Notably increased parietal activation (BA39 and BA7), which have been implicated in attention modulation, goal-directed behaviours and semantic processing (Silk et al., 2005). The increased activation in the parietofrontal network is interpreted to indicate that ASD individuals rely more heavily on this network for attentional and integrative purposes, suggesting an alternate compensatory strategy (Silk et al., 2005; Cohen, 2009). Similarly, there is research reporting that when task demands allow for more cognitive or explicit processing, ASD individuals tend to show more typical-like brain activation in regions that are implicated in emotion processing (Critchley et al., 2000; Piggot et al., 2004; Wang et al., 2004). For example, Wang, Lee, Sigman, and Dapretto (2007) reported that when explicitly instructed to attend to certain social cues, ASD individuals showed increased activation of the inferior frontal cortex and temporal regions bilaterally compared to the task when they are not explicitly instructed to do so.

Understanding multi-sensory information becomes even more complicated when the information from different sources conflict with the other. A number of studies have sought to elucidate the neural underpinnings of audio-visual incongruence processing in the typical population (e.g., Carter and Van Veen, 2007; Haas et al., 2006; Kerns et al., 2004; Ochsner et al., 2009). This line of research has emphasized the significant role of the anterior cingulate

cortex (ACC) and the prefrontal regions, such as the dorsolateral prefrontal cortex (dlPFC) and posterior medial frontal cortex (pmFC) in detecting and resolving incongruence. There is research investigating congruence or incongruence processing in audiovisual emotional processing (e.g., Klasen et al., 2011; Müller et al., 2011), in which brain regions that respond more to incongruent than to congruent audiovisual stimulation would be presumed to be involved in a conflict process. For example, Watson et al. (2013) examined the neural correlates of audiovisual emotional incongruence processing and reported that incongruent compared to congruent emotions evoked increased activity in the superior temporal region, implying the detection of mismatching information in this region.

In addition, Müller et al. (2011, 2014) distinguished between two types of conflicts: one is the incongruence of emotion-valence in which the emotion-valence in two sensory channels of audiovisual presentations are different; the other is the incongruence of emotion-presence in which one of the two sensory channels in an audio-visual stimulus has a neutral stimulus and the other channel has an emotional stimulus. By comparing emotional incongruence conditions to emotional congruence conditions, Müller et al. (2011) reported an incongruence effect of emotion-valence in the middle cingulate cortex, right superior frontal cortex, right supplementary motor area as well as the right temporoparietal junction, suggesting a role of these regions in bimodal conflict processing. By comparing emotion-presence conditions with emotion-absence conditions, this study identified greater activation in the left lingual gyrus, left middle temporal gyrus, right calcarine and the right superior temporal gyrus.

Despite this evidence in the neurotypical population, the neural correlates of conflict processing still remain a relatively open question in ASD. There is one study by Watanabe et al. (2012) that assessed the neural correlates when adults with and without ASD were processing

non-matching cues between verbal and nonverbal information and reported a significant decrease in brain activation in a number of regions in individuals with ASD, including the right inferior frontal gyrus (rIFG), bilateral anterior insula (AIC), ACC/ventral medial prefrontal cortex (vmPFC), and dorsal medial prefrontal cortex (dmPFC). Anomalies of brain activation were also observed in children/adolescents with ASD when dealing with conflict arising from different information sources. Specifically, Wang et al. (2007) reported a decrease in activation in the mPFC and right superior temporal gyrus (rSTG) in children/adolescents with ASD during an irony comprehension task, though the level of activation in mPFC was modulated by an explicit instruction phase. These results may suggest less automatically integrative reasoning about intent in the ASD population.

The heterogeneity of severity and the varying extent of the symptoms that individuals with ASD exhibit has led researchers to suggest that these disorders lie on a continuum of social-communication difficulties which extend into the general population (Baron-Cohen, 1995). The autism-spectrum quotient (AQ) was developed to measure the degree to which individuals with normal intelligence demonstrate traits of the core autistic phenotype (Baron-Cohen et al., 2001). There is evidence suggesting that autistic traits as measured by the AQ predict autistic-like behavioral performance, including decreased global processing (Grinter et al., 2009; Neufeld et al., 2020), difficulties in mental state reading from eyes (Baron-Cohen et al., 2001), and local-biased attention towards gaze orientation (Bayliss and Tipper, 2005). Additionally, von dem Hagen et al. (2011) showed an effect of autism spectrum traits on both the structure and function of posterior superior temporal sulcus, a region that is thought to be implicated in ASD regarding social cognition.

There is research reporting that individuals with significant autistic traits exhibited elevated levels of loneliness in relation to lack of social skills and understanding, relative to those with low levels of autistic traits (Austin, 2005; Jobe and White, 2007). It is possible that even mild traits of autism, if common, may lead to considerable social difficulties in real-life situations and cost for public health. Considering that some authors consider ASD to be at the extreme end of a quantitative variation of autistic traits (Constantino and Todd, 2003), investigating the incongruence effect in autism traits would enhance our understanding of audiovisual emotion processing with autistic traits and this could motivate further study in people with a diagnosis of ASD. Moreover, improving our understanding of the relationship between subclinical traits, neural function and multisensory processing of emotion in the general population will help to provide information about - potential behavioural and neurobiological correlates of autistic traits. Accordingly, the first aim of the present study was to examine the neural correlates by which the brain detects and resolves emotional conflict, and how these process change with autistic traits. Considering that an incongruence also arises from concurrent presentation of an emotional stimulus in one sensory channel and a neutral one in the other (Müller et al., 2011), we also investigated the brain correlates of incongruence regarding the presence and absence of emotion, by contrasting audiovisual stimulus where two sensory channels were both emotional (emotion-presence conditions) with audiovisual stimulus where an emotional cue in one channel and a neutral cue in the other (emotion-absence conditions). In line with previous findings on incongruence processing in ASD (e.g., Wang et al., 2007; Watanabe et al., 2012), we predicted that individuals with substantial autistic traits would present with reduced activation in regions that were engaged in incongruence processing (e.g., ACC or prefrontal regions).

2. Materials and methods

2.1 Participants

A total of 34 right-handed participants, as determined by the Edinburgh Handedness Inventory Questionnaire (Oldfield, 1971), were involved in this study. Written informed consent was obtained for each participant according to a protocol approved by the ethics committee of the College of Science and Engineering, University of Glasgow. All participants completed the Autism Spectrum Quotient (AQ) (Baron-Cohen et al., 2001), a common measure to assess autism spectrum traits in both the general population and the autism spectrum community. Individuals with a diagnosis of an autism spectrum disorder or clinically significant autistic symptoms ($AQ \geq 32$) were assigned to the high autism-spectrum quotient score (HAQ) group, while participants who scored at 18 or lower on the AQ were assigned to the low autism-spectrum quotient score (LAQ) group (Robertson and Simmons, 2013). This led that our HAQ group consisted of those who scored 31 or more on the AQ (High: 31-50) as one person with a formal diagnosis of Asperger Syndrome scored at 31 and our LAQ consisted of those who score at 18 or lower on the AQ (Low: 0-18). All participants were reported to have normal or corrected to normal visual and hearing acuity. Five participants were excluded from the analysis due to full scan incompleteness or run-time errors. Therefore, 16 individuals (11 females, 5 males) with HAQ and 13 individuals (10 females, 3 males) with LAQ were included in the final analysis, as shown in Table 1.

Table 1. Demographic characteristics of the HAQ group and LAQ group

Characteristic	HAQ(N=16)	LAQ(N=13)
Autism Quotient**	38.67 (5.73)	12.15(4.1)
Age	31.88 (10.29)	23.23(4.38)
Gender	11F, 5M	10F,3M

Notes. ** $p < .01$, * $p < .05$

2.2 Stimuli

The stimuli were face-voice pairs. The face information, selected from the Binghamton-Pittsburgh 4D Spontaneous Facial Expression Database (Zhang et al., 2014), consisted of facial behaviors of 5 males and 5 females, each showing happy, neutral, and fearful expressions, resulting in 30 different facial behaviors. The voice information consisted of 10 audible yawns, 10 laughs and 10 screams (5 males and 5 females for each), selected from Montreal Affective Voices (MAV; Belin et al., 2008). The pairing was done by presenting emotional voices (e.g., happy, fearful) during expressive facial movements in the clip, and presenting neutral voices during non-expressive facial movements in the middle of the clip, resulting in a total of nine types of emotion-congruent face-voice stimulus (con-FV) and emotion-incongruent face-voice stimulus (incon-FV): a happy face paired with a happy voice (H/H), a happy face paired with a fearful voice (H/F), a happy face paired with a neutral voice (H/N), a fearful face paired with a fearful voice (F/F), a fearful face paired with a happy voice (F/H), a fearful face paired with a neutral voice (F/N), a neutral face paired with a neutral voice (N/N), a neutral face paired with a happy voice (N/H), a neutral face paired with a fearful voice (N/F). In addition, the pairing was matched with regard to gender, in the sense that a female face was always paired with a female voice and vice versa. All face-voice pairs were 3s in duration with a frame rate of 25 frames per second and were presented on a black background. The voice information was presented binaurally via electrostatic earphones (NordicNeuroLab, Norway) with an average sound intensity of 80dB.

2.3 Procedure

Two functional runs were performed during a fMRI session. Each run consisted of 9 blocks of face-attend condition and 9 blocks of voice-attend condition presented in a pseudorandomized order using Presentation 14.9, designed by NeuroBehavioral Systems (NBS). Each block consisted of 5 face-voice pairs, lasting 16s. Each run began with 20 seconds of a black screen before the stimulus trials began and ended with 12 seconds of a black screen. In order to avoid any possible order effects, three different pseudo-random sequences were created, and each sequence was shown to a third of the participants (the third sequence was shown to 12 participants). Before the onset of each block, participants viewed a white color fixation cross on a black background. In order to ensure that participants were alert and attentive, on one-third of the trials the white color fixation cross was replaced by a red color fixation cross. Participants indicated with a button press when they saw a red color fixation cross. For the face-attend conditions, participants were asked to attend to the face and ignore the voice. For the voice-attend conditions, participants were asked to attend to the voice and ignore the face.

2.4 fMRI data acquisition

All fMRI data were acquired on a 3T Tim Trio MRI scanner (Siemens, Erlangen, Germany). Functional T2 weighted images were recorded with an EPI sequence (TR=2000msec, TE=30msec, interleaved, flip angle 90°, spatial resolution = 3.0 × 3.0 × 3.0 mm), using a 32-channel head coil. A whole brain anatomical scan was performed using a T1-weighted magnetization prepared rapid gradient echo imaging (MPRAGE) acquisition (192 contiguous 1mm axial slices, dimensions: 256mm*256mm, TR= 1900ms, TE=2.52ms, inversion time= 900ms, flip angle=9°), lasting 5 minutes, TR= 2300ms, TE = 2.91 ms, flip angle = 9° , spatial resolution 1.0 × 1.0 × 1.2 mm resolution).

2.5 Data analysis

Both fMRI data pre-processing and analysis were performed using Brainvoyager QX(2.8) (Brain Innovation). Structural data were transformed into Talairach space (Talairach and Tournoux, 1988). Functional images were slice scan-time corrected, motion corrected and temporally filtered at 2 cycles across the time-course. Spatial smoothing with a Gaussian kernel of 6 mm FWHM was also applied. After this, the functional runs were aligned to the anatomical data sets, transformed into a Talairach space and converted into 4D volume time-course data.

A random-effects general linear model (RFX GLM) was used to calculate the first-level time-series analysis on the z-normalized BOLD signal for each subject. Each experimental condition was separately modelled in the first-level (single-subject) time-series analysis. Three rest blocks were also modeled in the first-level time-series analysis and regarded as condition of no interest. Head motion correction parameters were included in the GLM model as nuisance predictors. These individual first-level time-series were then entered into second-level random effects analysis to construct statistical parametric maps at the group level. Specifically, in order to investigate the incongruence effect of emotion-valence, a 2 (group: LAQ, HAQ) \times 2 (attend: face-attend, voice-attend) \times 2 (inconsistency: emotion-valence incongruent, emotion-valence congruent) ANOVA was undertaken, with factors of attend and inconsistency as within-subject variables, and factor of group as the between-group variable. Similarly, a 2 (group: LAQ, HAQ) \times 2 (attend: face-attend, voice-attend) \times 2 (inconsistency: emotion-presence, emotion-absence) ANOVA was performed to determine effect of group, attend and emotion-presence inconsistency on percentage BOLD signal change per voxel, with factors of attend and emotion-presence inconsistency as within-subject variables, and factor of group as the between-group variable. Significant results of the ANOVA were followed with post-hoc t-test by extracting beta values

from the activated brain regions. All created F -maps were thresholded at a significance level $P < 0.001$ (uncorrected). The resulting statistical maps were then submitted to a whole-brain correction criterion based on the estimate of the map's spatial smoothness and on an iterative procedure (Monte Carlo simulation) for estimating cluster-level false-positive rate. After 1000 iterations of a Monte Carlo simulation, the minimum cluster size threshold that yielded a cluster-level false-positive rate of 5% was applied to the statistical maps.

3. Results

3.1 Incongruence effect of emotion-valence

Results for the main effects of group, emotion-valence inconsistency and attend, as well as their interactions, obtained in the whole-brain analysis are shown in Table 2.

Main effects

A significant main effect of group was observed in the right inferior parietal lobule (IPL), with a cluster size threshold of $p < 0.001$, 9 contiguous functional voxels (Fig.1A and Table 2a). Activity extracted from this cluster was significantly greater for the HAQ group than the LAQ group ($t=4.42$, $p=0.0001$).

A main effect of emotion-valence inconsistency was observed in several brain regions. These brain regions centered on the bilateral inferior frontal gyrus (IFG), right cuneus, superior occipital gyrus (SOG), inferior parietal lobule (IPL), precuneus, anterior cingulate, the left medial frontal gyrus (MFG) and the superior temporal gyrus (STG) ($p < 0.001$, 24 contiguous functional voxels) (Table 2b). Examination of the direction of the effect indicated that the greater incongruence effect of emotion-valence, as defined by emotion-valence incongruence conditions

contrasted with emotion-valence congruence conditions, was found in regions of the bilateral IFG (right IFG [$t=3.89$, $p=0.0003$]; left IFG [$t=3.11$, $p=0.003$]), right cuneus ($t=5.48$, $p<0.0001$), left MFG ($t=3.35$, $p=0.001$) and STG ($t=4.82$, $p<0.0001$) (Fig. 2.A). The regions of right SOG ($t=3.77$, $p=0.0004$), IPL ($t=3.90$, $p=0.0003$), precuneus ($t=5.17$, $p<0.0001$), and anterior cingulate ($t=4.09$, $p=0.0001$) showed greater response to the emotion-valence congruence conditions than the emotion-valence incongruence conditions.

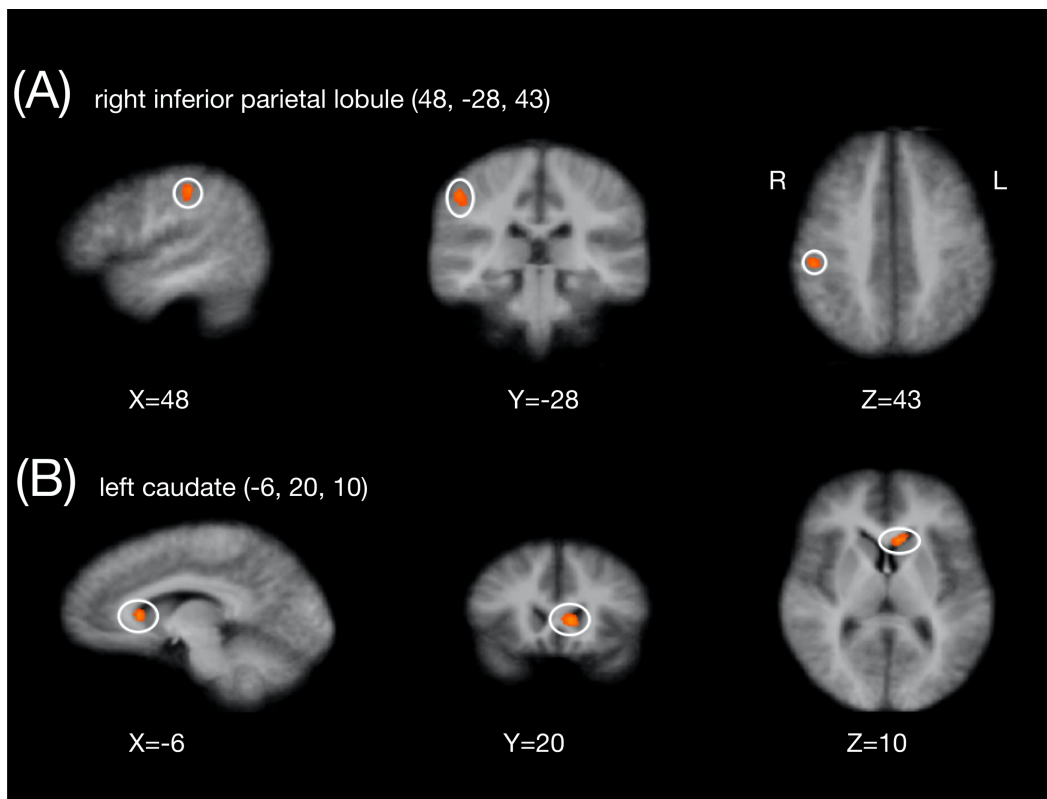


Fig.1. Main effect of group and interaction effect of group * attend* emotion-valence inconsistency. (A) Brain activation in the right inferior parietal lobule for the main effect of group ($p<0.001$ uncorrected, with a cluster size threshold of 9 contiguous functional voxels). (B) Brain activation in the left caudate for the three-way interaction of group * attend *emotion-valence inconsistency ($p<0.001$, with a cluster size threshold of 13 contiguous functional voxels).

Interaction effects of attend and/or emotion-valence inconsistency by group

There was no two-way interaction effect of group-by-emotion-valence inconsistency. However, the two way interaction effect of group-by-attend activated the regions of right medial frontal gyrus (MFG) and the left anterior cingulate (ACC) although this did not satisfy the cluster-size threshold criteria ($p < 0.001$, 11 contiguous functional voxels). Notably, these two regions were the same clusters as revealed from the ANOVA for the incongruence effect of emotion-presence. Later post-hoc t-tests did not find significant differences between the two groups in both face-attend conditions and voice-attend conditions.

The three-way interaction of group*attend*emotion-valence inconsistency was observed in the left caudate ($p < 0.001$, with a cluster size threshold of 13 contiguous functional voxels), indicating that the effect of emotion-valence inconsistency differentially influenced face-attend and voice-attend conditions for the two groups (Fig.1B and Table 2e). In order to understand this effect, the cluster's beta values were extracted and further analyzed for two levels of attend (face-attend, voice-attend) at the emotion-valence incongruence conditions as well as emotion-valence congruence conditions. Of most interest in this analysis was the incongruence effect of emotion-valence (defined by emotion-valence incongruence contrasted with emotion-valence congruence) between groups. This produced the result that, when instructed to attend to the visual track, the LAQ group exhibited greater caudate response to the incongruence effect of emotion-valence, compared to the HAQ group ($t = 2.91$, $p = 0.007$). When instructed to attend to the auditory track, the HAQ group exhibited greater response to the incongruence effect of emotion-valence, as compared to the LAQ group ($t = 3.19$, $p = 0.004$).

Table 2. The incongruence effect of emotion-valence. Clusters showing significant activations for: (a) main effect of group, (b) main effect of emotion-valence inconsistency, (c) group by attend interaction, (d) group by emotion-valence inconsistency interaction (e) group *attend * emotion-valence inconsistency interaction

Hemisphere	Anatomical region	Talairach coordinate of peak voxel (x, y, z)	BA	F-value	P-value	No. of voxels
(a) Group (LAQ, HAQ)						
R	Inferior Parietal Lobule	(48,-28,43)	40	25.93	<0.0001	311
(b) Emotion-valence inconsistency (incongruence, congruence)						
R	Cuneus	(20, -96, 3)	18	211.695969	<0.00001	91,080
R	Inferior Frontal Gyrus	(44, 28, 6)	13	43.495922	<0.00001	8419
R	Superior Occipital Gyrus	(35, -77, 27)	19	45.206665	<0.00001	2232
R	Inferior Parietal Lobule	(38, -44, 39)	40	36.651894	<0.00001	2108
R	Precuneus	(14, -65, 18)	31	113.35891	<0.00001	125,019
R	Anterior Cingulate	(11, 43, 9)	32	83.832161	<0.00001	54,671
L	Medial Frontal Gyrus	(-10, 52, 39)	6	37.367836	<0.00001	2641
L	Superior Temporal Gyrus	(-43, -29, 6)	41	222.232941	<0.00001	24,715
L	Inferior Frontal Gyrus	(-40, 31, -3)	47	45.645294	<0.00001	8919
(c) Group (LAQ, HAQ) * attend (face-attend, voice-attend)						
R	Medial Frontal Gyrus	(12, 41, -9)	10	21.45	0.0002	185
L	Anterior Cingulate	(-15, 44, -2)	32	19.32	0.0001	159
(d) Group (LAQ, HAQ) *emotion-valence inconsistency (incongruence, congruence)						
non-significant						
(e) Group (LAQ, HAQ) * attend (face-attend, voice-attend) * emotion-valence inconsistency (congruence, incongruence)						
L	Caudate	(-6, 20, 10)	-	32.53	<0.00001	388

x, y, z, Talairach coordinates of voxel with peak-activation. L=left hemisphere, R= right hemisphere. BA = Brodmann's area. All the reported regions were thresholded at $p < 0.001$, (uncorrected).

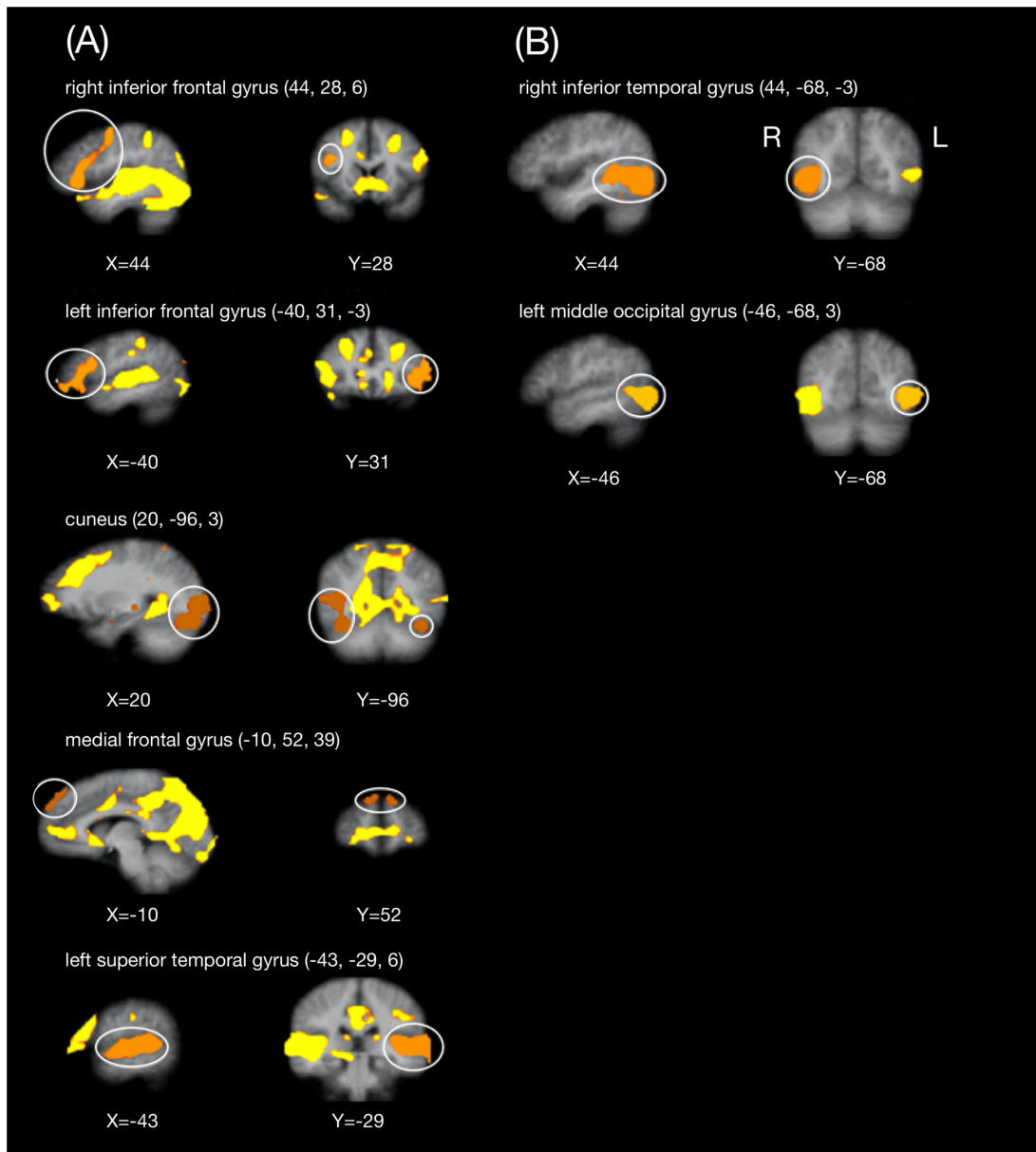


Fig 2. Sagittal (SAG) and Coronal (COR) views showing significant greater BOLD signal for main effect of emotion-valence inconsistency and main effect of emotion-presence inconsistency at Talairach coordinates. Significant clusters are circled in white. (A) Brain regions showed greater response for the emotion-valence incongruence contrasted with emotion-valence congruence conditions (incongruence effect of emotion-valence) ($p < 0.001$ uncorrected, with a cluster size threshold of 24 contiguous functional voxels). (B) Brain regions showed greater response for the emotion-presence contrasted with emotion-absence conditions (incongruence of emotion-presence) ($p < 0.001$ uncorrected, with a cluster size threshold of 23 contiguous functional voxels).

3.2 Incongruence effect of emotion-presence

Results for the main effects of group, attend, and emotion-presence inconsistency, as well as their interactions, obtained in the whole-brain analysis are shown in Table 3.

Main effects

The ANOVA (group * attend* emotion-presence inconsistency) resulted in a main effect of group (LAQ, HAQ) in the right IPL ($p < 0.001$, with a spatial extent 7 functional voxels) (Fig.3A and Table 3a). However, this region did not satisfy the cluster-size threshold criteria ($p < 0.001$, 9 contiguous functional voxels). For the sake of completeness, post hoc analysis on the beta values for this region was further conducted between the two groups over all conditions. This suggested that the HAQ group showed increased response to the overall conditions compared to the LAQ group ($t = 4.40$, $p = 0.0002$).

A significant main effect of emotion-presence inconsistency was observed in two large clusters ($p < 0.001$, with a cluster size threshold of 23 contiguous functional voxels) (Fig.2B and Table 3b). One cluster was centered on the right inferior temporal gyrus (ITG), extending to middle temporal gyrus (MTG), superior temporal gyrus (STG) and middle occipital gyrus (MOG). The other cluster was centered on the left middle occipital gyrus (MOG), extending to the middle temporal gyrus (MTG) and the inferior temporal gyrus (ITG). Post-hoc t-tests were conducted using the beta values for emotion-presence conditions and emotion-absence conditions at the peak coordinates of the two above clusters ($x = 44$, $y = -68$, $z = -3$; $x = -46$, $y = -68$, $z = 3$). Results suggested increased activation for the emotion-presence conditions where the paired face and voice in audio-visual presentations were both emotional (H/H, H/F, F/F, F/H), compared with emotion-absence conditions where one of the stimuli (face or voice) was neutral (H/N, F/N, N/H, N/F) ($t = 3.09$, $p = 0.003$; $t = 2.60$, $p = 0.01$, respectively).

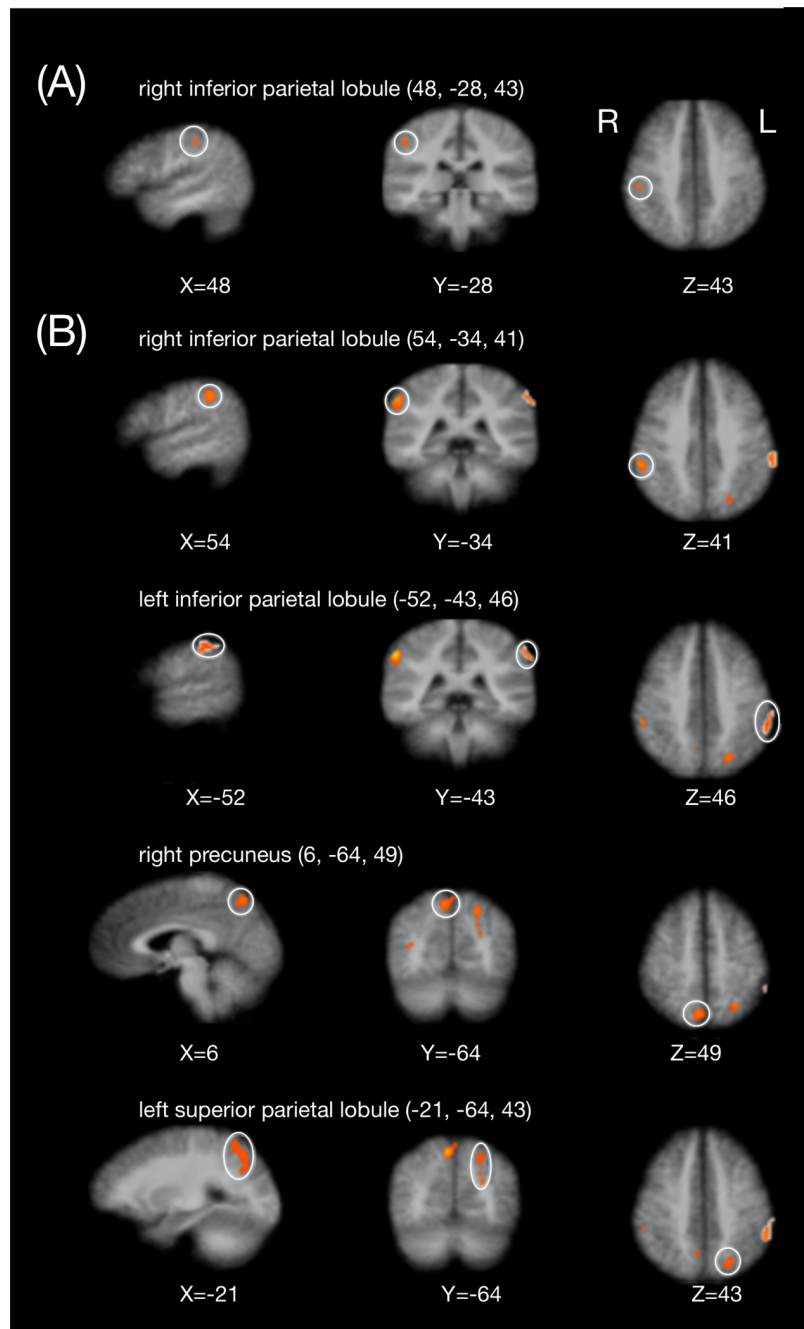


Fig3. Main effect of group and interaction effect of group * attend* emotion-presence inconsistency. (A) Brain activation in the right inferior parietal lobule for the main effect of group ($p < 0.001$ uncorrected, with a spatial extent 7 functional voxels). (B) Brain activation in regions of the bilateral inferior parietal lobule, right precuneus and left superior parietal lobule for the three-way interaction of group * attend * emotion-presence inconsistency ($p < 0.001$, with a cluster size threshold of 15 contiguous functional voxels).

Interaction effects of attend and/or emotion-presence inconsistency by group.

Testing for group difference, a group by emotion-presence inconsistency interaction was found in the left parahippocampal gyrus (PHG) ($p < 0.001$, clustered at 9 contiguous functional voxels) (Table 3c). Post-hoc analysis revealed that this interaction effect was driven by a greater response to the emotion-presence conditions in the HAQ group, compared to the LAQ group ($t = 2.33$, $p = 0.03$). Two groups showed no significant difference in response to the emotion-absence conditions ($t = -0.25$, $p = 0.81$).

There were two clusters in the right medial frontal gyrus (MFG) and the left anterior cingulate (AC) showing significant two-way interaction of group-by-attend ($p < 0.001$, with a cluster size threshold of 11 contiguous functional voxels). Follow-up analysis of the beta weights revealed no significant difference in BOLD activity between the HAQ and the LAQ groups in both the face-attend conditions as well as the voice-attend conditions in these two regions (all $p > 0.05$).

Finally, a three-way interaction of group* attend* emotion-presence inconsistency was found in a number of posterior brain regions, involving bilateral inferior parietal lobule (IPL), right precuneus and left superior parietal lobule (SPL) ($p < 0.001$, 15 contiguous functional voxels) (Fig.3B and Table 3e). Then post hoc analysis was performed on the beta values for two levels of attend (face-attend, voice-attend) in the emotion-presence conditions, as well as among those in the emotion-absence conditions at the peak coordinates of each significant cluster ($x = 54$, $y = -34$, $z = 41$; $x = 6$, $y = -64$, $z = 49$; $x = -21$, $y = -64$, $z = 43$; $x = -52$, $y = -43$, $z = 46$). We made these comparisons between groups as we were primarily interested in exploring the incongruence effect of emotion-presence (defined by emotion-presence conditions contrasted with emotion-absence conditions) between groups. These results suggest that, for the face-attend conditions, the HAQ group

exhibited greater activity in response to the incongruence effect of emotion-presence, compared to the LAQ group ($t=3.19$, $p=0.004$; $t=3.17$, $p=0.004$; $t=3.54$, $p=0.001$; $t=2.43$, $p=0.02$, respectively). However, for the voice-attend conditions, the LAQ group showed a greater incongruence effect of emotion-presence than the HAQ group ($t=2.18$, $p=0.04$; $t=2.35$, $p=0.03$; $t=2.65$, $p=0.01$; $t=3.55$, $p=0.001$, respectively).

Table 3. The incongruence effect of emotion-presence. Clusters showing significant activations for: (a) main effect of group, (b) main effect of emotion-presence, (c) group-by-attend interaction, (d) group-by-emotion-presence inconsistency interaction (e) group-by-emotion-presence inconsistency-attend interaction

Hemisphere	Anatomical region	Talairach coordinate of peak voxel (x, y, z)	BA	F-value	Uncorrected	No. of voxels
(a) Group (LAQ, HAQ)						
R	Inferior Parietal Lobule	(48,-28,43)	40	25.1	<0.0001	212
(b) Emotion-presence inconsistency (incongruence, congruence)						
R	Inferior Temporal Gyrus	(44, -68, -3)	37	107.7	<0.00001	16,703
L	Middle Occipital Gyrus	(-46, -68, 3)	37	66.6	<0.00001	6717
(c) Group (LAQ, HAQ) * attend (face-attend, voice-attend)						
R	Medial Frontal Gyrus	(12, 41, -9)	10	21.45	<0.001	319
L	Anterior Cingulate	(-15, 44, -2)	32	19.32	<0.001	339
(d) Group (LAQ, HAQ) * emotion-presence inconsistency (emotion-presence, emotion-absence)						
L	Parahippocampal Gyrus	(-15, -40, 7)	30	19.33	<0.001	305
(e) Group (LAQ, HAQ) *attend (face-attend, voice-attend) * emotion-presence inconsistency (emotion-presence, emotion-absence)						
R	Inferior Parietal Lobule	(54, -34, 41)	40	26.34	<0.0001	485
R	Precuneus	(6, -64, 49)	7	26.65	<0.0001	806
L	Superior Parietal Lobule	(-21, -64, 43)	7	19.93	0.0001	1237
L	Inferior Parietal Lobule	(-52, -43, 46)	40	21.87	<0.0001	1298

x, y, z, Talairach coordinates of voxel with peak-activation. L=left hemisphere, R= right hemisphere. BA = Brodmann's area. All the reported regions were thresholded at $p<0.001$, (uncorrected).

4. Discussion

Our study investigated how individuals with substantial autistic traits (HAQ) and individuals with few autistic traits (LAQ) respond to different types of incongruence during cross-modal processing: 1) incongruence of emotion-valence, where cues from two sensory channels were either congruent or incongruent to their emotion-valence, 2) incongruence of emotion-presence, where the paired face and voice in audio-visual presentations were either both emotional (emotion-presence) or one of the stimuli (voice or face) was neutral (emotion-absence). To do so, whole brain fMRI data were collected when 16 adults with high autism-spectrum quotient score and 13 adults with low autism-spectrum quotient score were instructed to perceive emotional (happy, fearful) and neutral faces or voices while concurrently being presented with emotional (fearful, happy) voices or faces.

Group effects

Independent of the track participants were told to attend, we found a main effect of group on brain activation for contrasts involving both emotion-valence and emotion-presence, where the HAQ group showed greater response compared to the LAQ group in the right IPL. In typical individuals, the IPL has been associated with multiple functions including attentional selection and response inhibition (Cai et al., 2014; Kaping et al., 2011), action observation and representation (Bara et al., 2011), motor imagery (Thinnes-Elker et al., 2012; Tunik et al., 2007) and integrative sensorimotor processing (Fogassi and Luppino, 2005; Hickok and Poeppel, 2004). Important for interpreting our findings are recent results in typical individuals that have emphasized a domain-general role of the IPL in processing emotional information. For example, a lesion study by (Becker et al., 2012) reported that patients with complete bilateral amygdala damage showed intact recognition of fearful or negative facial expressions as well as increased

activation in the premotor cortex and the IPL to fearful faces. Using multi-site transcranial magnetic stimulation (TMS), Engelen et al. (2015) reported that stimulation of the IPL selectively led to enhanced fearful body processing selectively, providing the first evidence that IPL plays a causal role in emotional body processing. Further evidence for this was found by Mazzoni et al. (2017) who used a state-dependent TMS paradigm to demonstrate that the anterior intraparietal sulcus was involved in representing affective body movements. Moreover, increased functional coupling between the amygdala and IPL was reported in a study contrasting emotional faces with neutral faces, suggesting interaction between these two brain regions (Diano et al., 2017). Using continuous theta burst stimulation (cTBS)-fMRI, Engelen et al. (2018) reported increased amygdala activation during the observation of angry stimuli compared to neutral stimuli after stimulation of IPL/PMC, implying that the IPL communicates with amygdala to process the emotional content of actions.

Relating these findings about IPL to autism, our results found increased activity in the IPL for high ASD trait individuals compared to low ASD traits individuals while perceiving audiovisual emotional stimuli. Previous fMRI studies examining emotional processing in ASD reported atypical brain response in the IPL. For example, there is research reporting ASD-related abnormality in response to social (emotional) touch, including decreased neural reactivity to pleasant touch in IPL and social-emotional brain circuitry compared to typically developed controls (Kaiser et al., 2016), and reduced connectivity between the parietal operculum and major pain networks as well as increased connectivity between the semantic and limbic networks (Masson et al., 2020). Related to the processing of human biological, differences in IPL activation have been found. For example, Freitag et al. (2008) found that compared to the typically developed individuals, autistic individuals showed less brain activation for processing

biological motions in the IPL, IPS, STSp, or FG. Further, despite comparable behavioural performance on a biological motion perception task, McKay et al. (2012) identified that the typically developed group engaged the IPL in the network processing biological motion but the ASD group did not. In addition to these results, several neuroimaging studies observed more activation in the parietal regions in individuals with ASD compared to typically developed controls (Dapretto et al., 2006; Jack and Morris, 2014; Su et al., 2020). For example, during imitating and observing emotional facial expressions, unlike typically developed controls who showed more bilateral IFG activation, children with ASD showed greater activity in the left parietal and right visual association cortices but no activity in the IFG (Dapretto et al., 2006). Increased IPL in the HAQ group in present study extends these past literature of atypical activation in the IPL to emotional audiovisual displays.

Although the significance of IPL in responding to emotional salient emotional information in the environment is well established, it is potentially significant that IPL is linked to the attentional network, especially in the reorienting or shifting of attention (Singh-Curry and Husain, 2009). This region has also been implicated in cognitive control and mediating the allocation of attention to task-relevant information (Ciaramelli et al., 2008). Apart from communicating emotional information in the stimuli, perception of emotions also guide one's behavioural responses. This has led to the suggestion that perception of emotions may not depend on roles of attention or social cognition in IPL in isolation, but on their overlapping functions of attention and social cognition. As such, current results of increased activity in the IPL could reflect a distinct perceptual strategy to form audiovisual emotional representations in individuals with high ASD traits, which may affect their social functioning and communication ability.

Incongruence effect of emotion-valence

In this section, the main effect of emotion-valence inconsistency, and significant interaction towards group * attention * emotion-valence inconsistency are discussed. Across both the LAQ and HAQ groups, a main effect of emotion-valence inconsistency obtained by contrasting emotion-valence incongruence conditions with emotion-valence congruence conditions led to activity primarily in fronto-temporal brain regions, including the bilateral IFG, MFG and STG. These brain regions have previously been associated with processing of incongruent information in typical populations (Mitchell, 2006; Watson et al., 2013; Wittfoth et al., 2010) and for the remainder of the paragraph we discuss these results from the typical population in more detail. The IFG has been suggested as a region that processes both cognitive and emotional conflict between stimulus dimensions (Aron et al., 2014; Li et al., 2020). Activity in IFG was reported in a number of studies of typical populations for interference monitoring and conflict resolution (Blasi et al., 2006; Dolcos and McCarthy, 2006). Mitchell (2006) examined the processing of emotional prosody in a conflicting semantic context with results suggesting a role of left IFG together with the bilateral STG and MTG and the basal ganglia in decoding incongruence. Our findings of the involvement of bilateral STG are also consistent with Watson et al. (2013) who reported that incongruent compared to congruent conditions led to increased activation in the STG, suggesting a more general role of the STG in multimodal perception (Beauchamp, 2005; Miller and D'Esposito, 2005) and in the recognition and evaluation of socioemotional information (Kaiser et al., 2011; Scharpf et al., 2010). Apart from these findings, previous research has implicated regions in the MFG in response to conflict (Aarts et al., 2009; Woodward et al., 2006, 2008). Therefore, our results are consistent with a role of frontal-

temporal regions (including MFG, STG, and IFG) during emotion-valence incongruence processing over both LAQ and HAQ groups.

Importantly, a group * attend * emotion-valence inconsistency interaction was found in the left caudate. The caudate is part of the ventral striatum, which is an integral part of the fronto-striatal circuitry that is implicated in perceptual choices, selective attention (Herman et al., 2020), and executive cognitive functions (Grahn et al., 2008). Our results further elucidated that depending on whether participants were instructed to attend to the face or voice, different patterns of the incongruence effect of emotion-valence were found between the two groups. Specifically, a greater incongruence effect of emotion-valence occurred when instructed to attend to the auditory track for the HAQ group. Conversely, the LAQ group revealed a greater incongruence effect of emotion-valence when they were instructed to attend to the face. It seems possible that evaluation of emotional prosody with conflicting emotional faces was more attentionally demanding in the HAQ group, while for the LAQ group, evaluation of emotional faces with conflicting voice prosody was more attentionally demanding. These results, together with the previous findings, suggest that the two groups may adopt different strategies when decoding crossmodal emotional incongruence.

Incongruence effect of emotion-presence

In this section, the main effect of emotion-presence inconsistency, and the significant interactions of group * emotion-presence inconsistency, and group * attend * emotion-presence inconsistency are discussed. Across both the LAQ and HAQ groups, the main effect of emotion-presence inconsistency obtained by contrasting emotion-presence conditions with emotion-absence conditions led to activity in a set of temporal and occipital regions in both hemispheres, including the MOG, extending to ITG, MTG and STG. Our current results of emotion-presence

incongruence effect recruited similar brain network as a previous fMRI study that demonstrated increased activation in bilateral occipital and temporal cortex with respect to the emotion-presence incongruence effect (Müller et al., 2011). Regions in the temporal cortex, involving MTG and STG are well documented to be implicated in emotional and social information processing in typical populations, such as human voice (Belin et al., 2000), audiovisual speech (Wright et al., 2003), multimodal emotional integration or more generally multisensory processing (Kreifelts et al., 2007; Park et al., 2010; Peelen et al., 2010). Previous studies suggested that temporal cortex (e.g, MTG, STG) showed a heightened sensitivity to emotional cues, with a more pronounced response to emotional than neutral stimulation (Kreifelts et al., 2007). Using multivoxel pattern analysis, Peelen et al. (2010) suggested two regions of medial prefrontal cortex and STS in containing modality-independent, abstract representations of emotion, implying a key role of these regions in perceiving others' emotions. Via direct comparisons between emotion-presence conditions and emotion-absence conditions, our results showed that these occipital and temporal areas of MOG, ITG, MTG and STG maybe engaged in an emotion-sensitive network, which is specifically responsible for emotion-driven features during the processing of audiovisual emotional displays over both LAQ and HAQ groups.

In addition to the main effect of the emotion-presence inconsistency, a two-way interaction of group * emotion-presence inconsistency showed that the HAQ group demonstrated greater activity in the PHG in response to the emotion-presence conditions compared to the LAQ group. The PHG is known to play a critical role in the processing of information with socioemotional valence (Li et al., 2014), emotional memory encoding, and episodic memory retrieval (Murty et al., 2010; Nielson et al., 2015). Regarding emotional memory, the PHG has been implicated in processing negative emotional valence, suggesting that it may have a specific role in emotional

memory enhancements for scenes and locations (Murty et al., 2010). As such, increased activity of the PHG in the HAQ group with the current results may therefore reflect emotional memory enhancements, or the retrieval of socioemotional memories, or a combination of both in individuals with high autistic traits. If this is the case, such an interpretation would fit the previously suggested interpretation of our findings in which the increased IPL activity over all conditions in the HAQ group may be driven by overlapping functions of emotion and attention. Therefore, our results of increased activity in IPL and the PHG in individuals with HAQ may also be explained by the allocation of attention to salient emotional information (Rubia et al., 2010). This could suggest the possibility of different neural mechanisms underlying audiovisual emotional processing in individuals with high ASD traits, which would lead to more effortful and analytic processing of stimulus features, task-relevance or salience during the processing of audiovisual emotional displays. Given the connections of IPL and PHG with regions across temporal, parietal, and frontal cortices, it would be interesting in future studies to investigate how these two regions interact with other regions in the cortex and their roles in audiovisual emotional perception in individuals with high autistic traits or an ASD diagnosis.

The evaluation of emotional information depends upon not only emotional processing, but also executive control (Hart et al., 2010). As such, it is possible that the experimental instructions of which sensory track to attend could have modulated emotion-absence processing between the two groups. Such an explanation is relevant to the group * attend * emotion-presence inconsistency interaction we found in the bilateral IPL, right precuneus and left SPL. This suggests a different pattern of brain activation regarding an emotion-presence incongruence effect for the two groups of individuals, depending on which sensory track they were instructed to attend (face or voice). Specifically, for the HAQ group, a greater incongruence effect of

emotion-presence was only observed when attending to the visual track. However, for the LAQ group, a greater emotion-presence incongruence effect was found when participants were instructed to attend to the auditory modality. Regions in the parietal cortices (including the precuneus and parietal lobules) serve as an important part of dorsal visual stream and are implicated in visuospatial processing (Newman-Norlund et al., 2007) and visual attention shifts (Corbetta and Shulman, 2002). Thus, greater engagement of parietal cortices in face-to-be attended conditions may reflect a visual processing advantage in the HAQ group. Indeed, previous research provide evidence of enhanced visual processing in individuals with ASD in a range of tasks, such as faces, objects and words (for a review, see Samson et al., 2012). Several studies have observed deficits in multisensory facilitation in ASD. It has been shown that individuals with ASD tend to benefit less from the addition of auditory information to a visual search task (Collignon et al., 2013). Likewise, Charbonneau et al. (2013) suggested that individuals with ASD did not benefit as much from crossmodal emotion expressions as did typical populations. Schelinski et al. (2014, 2017) suggested both the voice identity and speech recognition are affected in individuals with ASD. There is also evidence of altered audiovisual modulation in ASD (Jao Keehn et al., 2017). Results for the auditory processing in this study reported less activity in the visual cortex for the typical populations, but greater activity in the ASD group, and linked these findings to autism symptomatology. Our present study examined the incongruence effect of emotion-presence in face-attend conditions and voice-attend conditions, this way we could explore the modulation of the to-be-attended sensory track (face or voice) in the processing of audiovisual emotional displays for the two groups. Our results suggest subtle processing differences in the parietal regions for the incongruence effect of emotion-presence in the two groups. Such an interpretation would fit the previously suggested

interpretation of a greater emotion-valence incongruence effect in face-attend conditions in the HAQ group. However, future work is needed to determine whether these differences in the HAQ group could be found in individuals with an ASD diagnosis.

There are potential limitations to the current study that need to be considered. First, the stimuli used in the current study combined dynamic facial behaviors and vocalisation. These facial behaviors indicated moment-to-moment changes in emotional states, providing a good approximation of naturalistic interaction. However, the stimuli may still have evoked some unnatural feelings in the participants due to being out of sync between facial behaviours and vocalisations. Future studies should attempt to address this limitation. Another possible concern is that the current study did not include an instruction-free condition, which prohibited us from making comparisons of regional brain activity between an instruction-free condition and the face-attend/voice-attend condition. Future investigations could take this into consideration and include an instruction-free task in the experimental design, in order to clarify the congruence effect in a more complete way. Additionally, this study was also limited by the small sample size, making it difficult to generalise the results to the wider population. In addition, although both groups were adult, the age difference between the groups might have contributed to the difference, with the older HAQ group potentially having more time to develop compensatory means to process multisensory emotional signals.

Another possible limitation is that the present study did not find an incongruence effect of emotion-valence at the ACC, despite previous evidence implicating this region in detecting and addressing conflicts (e.g., Carter and Van Veen, 2007; Egner et al., 2008; Ochsner et al., 2009; Wittfoth et al., 2010), especially when the incentives and/or task demands increased (e.g., stimuli changing from congruent to incongruent, or “response conflict”) (e.g., Smith et al., 2013).

The most likely explanation for this difference is the use of block design in the present study, in which stimuli of the same type (either congruent or incongruent) were presented for the whole block of time. This may have decreased the task demands or intensity of conflict, leading to the absence of involvement of the ACC. Future studies may incorporate an event-related design to provide additional evidence regarding this issue.

5. Conclusion

An important aspect of crossmodal interactions is the evaluation of crossmodal matching or conflict. This study examined two different types of emotional incongruence in audiovisual processing: incongruence of emotion-valence and incongruence of emotion-presence. The incongruence effect of emotion-valence was characterized by activation in fronto-temporal brain regions in both hemispheres involving the inferior frontal gyrus, cuneus, superior temporal gyrus, and middle frontal gyrus. The incongruence effect of emotion-presence was characterized by activation in a set of temporal and occipital regions in both hemispheres, including the middle occipital gyrus, middle temporal gyrus and inferior temporal gyrus. In addition, the present study identified greater recruitment of the right IPL in perceiving audio-visual emotional expressions in HAQ individuals, as compared to the LAQ individuals. Depending on whether face or the voice was-to-be attended, different patterns of emotional incongruence were found between the two groups. Specifically, the HAQ group tended to show more incidental processing to visual information whilst the LAQ group tended to show more incidental processing to auditory information during the crossmodal emotional incongruence decoding. These differences might be attributed to different attentional demands and different processing strategies between the two groups.

References

- Aarts, E., Roelofs, A., van Turenout, M., 2009. Attentional control of task and response in lateral and medial frontal cortex: brain activity and reaction time distributions. *Neuropsychologia* 47 (10), 2089–2099. <https://doi.org/10.1016/j.neuropsychologia.2009.03.019>.
- American Psychiatric Association, 2013. Diagnostic and statistical manual of mental disorders: DSM-5TM, 5th ed. American Psychiatric Publishing, Inc...
- Aron, A.R., Robbins, T.W., Poldrack, R.A., 2014. Inhibition and the right inferior frontal cortex: one decade on. *Trends Cognit. Sci.* 18 (4), 177–185. <https://doi.org/10.1016/j.tics.2013.12.003>.
- Austin, E.J., 2005. Personality correlates of the broader autism phenotype as assessed by the Autism Spectrum Quotient (AQ). *Pers. Individ. Differ.* 38 (2), 451–460. <https://doi.org/10.1016/j.paid.2004.04.022>.
- Bara, B.G., Ciaramidaro, A., Walter, H., Adenzato, M., 2011. Intentional minds : a philosophical analysis of intention tested through fMRI experiments involving people with schizophrenia , people with autism , and healthy individuals. *Front. Hum. Neurosci.* 5 (7). 1-11. <https://doi.org/10.3389/fnhum.2011.00007>.
- Baron-Cohen, S., 1995. *Mindblindness: An essay on autism and theory of mind*. MIT Press, Cambridge.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., Clubley, E., 2001. The Autism Spectrum Quotient : evidence from Asperger syndrome/high functioning autism, males and females, scientists and mathematicians. *J. Autism Dev. Disord.* 31 (1), 5–17. <https://doi.org/10.1023/A:1005653411471>.
- Bayliss, A.P., Tipper, S.P., 2005. Gaze and arrow cueing of attention reveals individual differences along the autism spectrum as a function of target context. *Br. J. Psychol.* 96 (1), 95–114. <https://doi.org/10.1348/000712604X15626>.
- Beauchamp, M.S., 2005. Statistical criteria in fMRI studies of multisensory integration. *Neuroinformatics* 3 (2), 93–113. <https://doi.org/10.1385/NI:3:2:093>.
- Becker, B., Mihov, Y., Scheele, D., Kendrick, K.M., Feinstein, J.S., Matusch, A., Aydin, M., Reich, H., Urbach, H., Oros-Peusquens, A.M., Shah, N.J., Kunz, W.S., Schlaepfer, T.E., Zilles, K., Maier, W., Hurlmann, R., 2012. Fear processing and social networking in the absence of a functional amygdala. *Biol. Psychiatr.* 72 (1), 70–77. <https://doi.org/10.1016/j.biopsych.2011.11.024>.
- Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B., 2000. Voice-selective areas in human auditory cortex. *Nature* 403 (6767), 309–312. <https://doi.org/10.1038/35002078>.
- Belin, P., Fillion-Bilodeau, S., Gosselin, F., 2008. The Montreal Affective Voices: a validated set of

- nonverbal affect bursts for research on auditory affective processing. *Behav. Res. Methods* 40 (2), 531–539. <https://doi.org/10.3758/BRM.40.2.531>.
- Blasi, G., Goldberg, T.E., Weickert, T., Das, S., Kohn, P., Zolnick, B., Callicott, J.H., Weinberger, D.R., Mattay, V.S., 2006. Brain regions underlying response inhibition and interference monitoring and suppression. *Eur. J. Neurosci.* 23 (6), 1658–1664. <https://doi.org/10.1111/j.1460-9568.2006.04680.x>.
- Bozikas, V.P., Kosmidis, M.H., Anezoulaki, D., Giannakou, M., Andreou, C., Karavatos, A., 2006. Impaired perception of affective prosody in schizophrenia. *J. Neuropsychiatry* 18 (1), 81–85. <https://doi.org/10.1176/jnp.18.1.81>.
- Cai, W., Cannistraci, C.J., Gore, J.C., Leung, H.C., 2014. Sensorimotor-independent prefrontal activity during response inhibition. *Hum. Brain Mapp.* 35 (5), 2119–2136. <https://doi.org/10.1002/hbm.22315>.
- Campanella, S., Belin, P., 2007. Integrating face and voice in person perception. *Trends Cognit. Sci.* 11 (12), 535–543. <https://doi.org/10.1016/j.tics.2007.10.001>.
- Carter, C.S., Van Veen, V., 2007. Anterior cingulate cortex and conflict detection: an update of theory and data. *Cognit. Affect Behav. Neurosci.* 7 (4), 367–379. <https://doi.org/10.3758/cabn.7.4.367>.
- Charbonneau, G., Bertone, A., Lepore, F., Nassim, M., Lasseonde, M., Mottron, L., Collignon, O., 2013. Multilevel alterations in the processing of audio-visual emotion expressions in autism spectrum disorders. *Neuropsychologia* 51 (5), 1002–1010. <https://doi.org/10.1016/j.neuropsychologia.2013.02.009>.
- Ciaramelli, E., Grady, C.L., Moscovitch, M., 2008. Top-down and bottom-up attention to memory: a hypothesis (AtoM) on the role of the posterior parietal cortex in memory retrieval. *Neuropsychologia* 46 (7), 1828–1851. <https://doi.org/10.1016/j.neuropsychologia.2008.03.022>.
- Cohen, Y.E., 2009. Multimodal activity in the parietal cortex. *Hear. Res.* 258 (1–2), 100–105. <https://doi.org/10.1016/j.heares.2009.01.011>.
- Collignon, O., Charbonneau, G., Peters, F., Nassim, M., Lasseonde, M., Lepore, F., Mottron, L., Bertone, A., 2013. Reduced multisensory facilitation in persons with autism. *Cortex* 49 (6), 1704–1710. <https://doi.org/10.1016/j.cortex.2012.06.001>.
- Constantino, J.N., Todd, R.D., 2003. Autistic traits in the general population: a twin study. *Arch. Gen. Psychiatr.* 60 (5), 524–530. <https://doi.org/10.1001/archpsyc.60.5.524>.
- Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. <https://doi.org/10.1038/nrn755>.
- Critchley, H.D., Daly, E.M., Bullmore, E.T., Williams, S.C.R., Amelvoort, T.V., Robertson, D.M., Rowe, A., Phillips, M., Mcalonan, G., Howlin, P., Murphy, D.G.M., 2000. The functional neuroanatomy of social behaviour Changes in cerebral blood flow when people with autistic disorder process facial

- expressions. *Brain* 123 (11), 2203–2212. <https://doi.org/10.1093/brain/123.11.2203>.
- Dapretto, M., Davies, M.S., Pfeifer, J.H., Scott, A.A., Sigman, M., Bookheimer, S.Y., Iacoboni, M., 2006. Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nat. Neurosci.* 9 (1), 28–30. <https://doi.org/10.1038/nm1611>.
- Dawson, G., Meltzoff, A.N., Osterling, J., Rinaldi, J., Brown, E., 1998. Children with autism fail to orient to naturally occurring social stimuli. *J. Autism Dev. Disord.* 28 (6), 479–485. <https://doi.org/10.1023/A:1026043926488>.
- Diano, M., Tamietto, M., Celeghin, A., Weiskrantz, L., Tatu, M.K., Bagnis, A., Duca, S., Geminiani, G., Cauda, F., Costa, T., 2017. Dynamic changes in amygdala psychophysiological connectivity reveal distinct neural networks for facial expressions of basic emotions. *Sci. Rep.* 7 (45260), 1–13. <https://doi.org/10.1038/srep45260>.
- Dolcos, F., McCarthy, G., 2006. Brain systems mediating cognitive interference by emotional distraction. *J. Neurosci.* 26 (7), 2072–2079. <https://doi.org/10.1523/JNEUROSCI.5042-05.2006>.
- Doyle-thomas, K.A.R., Goldberg, J., Szatmari, P., Hall, G.B.C., 2013. Neurofunctional underpinnings of audiovisual emotion processing in teens with autism spectrum disorders. *Front. Psychiatr.* 4 (48), 1–9. <https://doi.org/10.3389/fpsy.2013.00048>.
- Egner, T., Etkin, A., Gale, S., Hirsch, J., 2008. Dissociable neural systems resolve conflict from emotional versus nonemotional distracters. *Cerebr. Cortex* 18 (6), 1475–1484. <https://doi.org/10.1093/cercor/bhm179>.
- Engelen, T., de Graaf, T.A., Sack, A.T., de Gelder, B., 2015. A causal role for inferior parietal lobule in emotion body perception. *Cortex* 73, 195–202. <https://doi.org/10.1016/j.cortex.2015.08.013>.
- Engelen, T., Zhan, M., Sack, A.T., de Gelder, B., 2018. Dynamic interactions between emotion perception and action preparation for reacting to social threat :a combined cTBS-fMRI study. *ENeuro* 5 (3). <https://doi.org/10.1523/ENEURO.0408-17.2018>.
- Fogassi, L., Luppino, G., 2005. Motor functions of the parietal lobe. *Curr. Opin. Neurobiol.* 15 (6), 626–632. <https://doi.org/10.1016/j.conb.2005.10.015>.
- Freitag, C.M., Konrad, C., Häberlen, M., Kleser, C., von Gontard, A., Reith, W., Troje, N.F., Krick, C., 2008. Perception of biological motion in autism spectrum disorders. *Neuropsychologia* 46 (5), 1480–1494. <https://doi.org/10.1016/j.neuropsychologia.2007.12.025>.
- Gervais, H., Belin, P., Boddaert, N., Leboyer, M., Coez, A., Sfaello, I., Barthélémy, C., Brunelle, F., Samson, Y., Zilbovicius, M., 2004. Abnormal cortical voice processing in autism. *Nat. Neurosci.* 7 (8), 801–802. <https://doi.org/10.1038/nm1291>.
- Grahn, J.A., Parkinson, J.A., Owen, A.M., 2008. The cognitive functions of the caudate nucleus. *Prog.*

- Neurobiol. 86 (3), 141–155. <https://doi.org/10.1016/j.pneurobio.2008.09.004>.
- Grinter, E.J., Maybery, M.T., Van Beek, P.L., Pellicano, E., Badcock, J.C., Badcock, D.R., 2009. Global visual processing and self-rated autistic-like traits. *J. Autism Dev. Disord.* 39 (9), 1278–1290. <https://doi.org/10.1007/s10803-009-0740-5>.
- Haas, B.W., Omura, K., Constable, R.T., Canli, T., 2006. Interference produced by emotional conflict associated with anterior cingulate activation. *Cognit. Affect Behav. Neurosci.* 6 (2), 152–156. <https://doi.org/10.3758/CABN.6.2.152>.
- Hall, G.B.C., Szechtman, H., Nahmias, C., 2003. Enhanced salience and emotion recognition in Autism: a PET study. *Am. J. Psychiatr.* 160 (8), 1439–1441. <https://doi.org/10.1176/appi.ajp.160.8.1439>.
- Happé, F., Ehlers, S., Fletcher, P., Frith, U., Johansson, M., Gillberg, C., Dolan, R., Frackowiak, R., Frith, C., 1996. ‘Theory of mind’ in the brain. Evidence from a PET scan study of Asperger syndrome. *Neuroreport* 8 (1), 197–201. <https://doi.org/10.1097/00001756-199612200-00040>.
- Hart, S.J., Green, S.R., Casp, M., Belger, A., 2010. Emotional priming effects during Stroop task performance. *Neuroimage* 49 (3), 2662–2670. <https://doi.org/10.1016/j.neuroimage.2009.10.076>.
- Herman, J.P., Arcizet, F., Krauzlis, R.J., 2020. Attention-related modulation of caudate neurons depends on superior colliculus activity. *ELife* 9. <https://doi.org/10.7554/ELIFE.53998>.
- Hickok, G., Poeppel, D., 2004. Dorsal and ventral streams : a framework for understanding aspects of the functional anatomy of language. *Cognition* 92 (1–2), 67–99. <https://doi.org/10.1016/j.cognition.2003.10.011>.
- Hobson, R.P., Oustont, J., Lee, A., 1989. Naming emotion in faces and voices : abilities and disabilities in autism and mental retardation. *Br. J. Dev. Psychol.* 7 (3), 237–250. <https://doi.org/10.1111/j.2044-835X.1989.tb00803.x>.
- Jack, A., Morris, J.P., 2014. Neocerebellar contributions to social perception in adolescents with autism spectrum disorder. *Dev. Cogn. Neurosci.* 10, 77–92. <https://doi.org/10.1016/j.dcn.2014.08.001>.
- Jao Keehn, R.J., Sanchez, S.S., Stewart, C.R., Zhao, W., Grenesko-Stevens, E.L., Keehn, B., Müller, R.A., 2017. Impaired downregulation of visual cortex during auditory processing is associated with autism symptomatology in children and adolescents with autism spectrum disorder. *Autism Res.* 10 (1), 130–143. <https://doi.org/10.1002/aur.1636>.
- Jobe, L.E., White, S.W., 2007. Loneliness , social relationships , and a broader autism phenotype in college students. *Pers. Individ. Differ.* 42 (8), 1479–1489. <https://doi.org/10.1016/j.paid.2006.10.021>.
- Kaiser, M.D., Shiffrar, M., Pelphrey, K.A., 2011. Socially tuned : brain responses differentiating human and animal motion. *Soc. Neurosci.* 7 (3), 301–310. <https://doi.org/10.1080/17470919.2011.614003>.
- Kaiser, M.D., Yang, D.Y., Voos, A.C., Bennett, R.H., Gordon, I., Pretzsch, C., Beam, D., Keifer, C.,

- Eilbott, J., Mcglone, F., Pelphrey, K.A., 2016. Brain mechanisms for processing affective (and nonaffective) touch are atypical in autism. *Cerebr. Cortex* 26 (6), 2705–2714. <https://doi.org/10.1093/cercor/bhv125>.
- Kaping, D., Vinck, M., Hutchison, R.M., Everling, S., Womelsdorf, T., 2011. Specific contributions of ventromedial, anterior cingulate, and lateral prefrontal cortex for attentional selection and stimulus valuation. *PLoS Biol.* 9 (12), e1001224. <https://doi.org/10.1371/journal.pbio.1001224>.
- Kerns, J.G., Cohen, J.D., MacDonald, A.W., Cho, R.Y., Stenger, V.A., Carter, C.S., 2004. Anterior cingulate conflict monitoring and adjustments in control. *Science* 303 (303), 1023–1026. <https://doi.org/10.1126/science.1089910>.
- Klaser, M., Kenworthy, C.A., Mathiak, K.A., Kircher, T.T.J., Mathiak, K., 2011. Supramodal representation of emotions. *J. Neurosci.* 31 (38), 13635–13643. <https://doi.org/10.1523/JNEUROSCI.2833-11.2011>.
- Klin, A., 1991. Young autistic children's listening preferences in regard to speech: a possible characterization of the symptom of social withdrawal. *J. Autism Dev. Disord.* 21 (1), 29–42. <https://doi.org/10.1007/BF02206995>.
- Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., Wildgruber, D., 2007. Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. *Neuroimage* 37 (4), 1445–1456. <https://doi.org/10.1016/j.neuroimage.2007.06.020>.
- Li, Y., Qiao, L., Sun, J., Wei, D., Li, W., Qiu, J., Zhang, Q., Shi, H., 2014. Gender-specific neuroanatomical basis of behavioral inhibition/approach systems (BIS/BAS) in a large sample of young adults: a voxel-based morphometric investigation. *Behav. Brain Res.* 274, 400–408. <https://doi.org/10.1016/j.bbr.2014.08.041>.
- Li, J., Yang, X., Zhou, F., Liu, C., Wei, Z., Xin, F., Daumann, B., Daumann, J., Kendrick, K.M., Becker, B., 2020. Modafinil enhances cognitive, but not emotional conflict processing via enhanced inferior frontal gyrus activation and its communication with the dorsomedial prefrontal cortex. *Neuropsychopharmacology* 45 (6), 1026–1033. <https://doi.org/10.1038/s41386-020-0625-z>.
- Loveland, K.A., Steinberg, J.L., Pearson, D.A., Mansour, R., Reddoch, S., 2008. Judgments of auditory-visual affective congruence in adolescents with and without autism: a pilot study of a new task using fMRI. *Percept. Mot. Skills* 107 (2), 557–575. <https://doi.org/10.2466/pms.107.2.557-575>.
- Masson, H.L., de Beeck, H.O., Boets, B., 2020. Reduced task-dependent modulation of functional network architecture for positive versus negative affective touch processing in autism spectrum disorders. *Neuroimage* 219, 117009. <https://doi.org/10.1016/j.neuroimage.2020.117009>.
- Mazzoni, N., Jacobs, C., Venuti, P., Silvanto, J., Cattaneo, L., 2017. State-dependent TMS reveals

- representation of affective body movements in the anterior intraparietal cortex. *J. Neurosci.* 37 (30), 7231–7239. <https://doi.org/10.1523/JNEUROSCI.0913-17.2017>.
- McKay, L.S., Simmons, D.R., McAleer, P., Marjoram, D., Piggot, J., Pollick, F.E., 2012. Do distinct atypical cortical networks process biological motion information in adults with Autism Spectrum Disorders?. *Neuroimage* 59 (2), 1524–1533. <https://doi.org/10.1016/j.neuroimage.2011.08.033>.
- Miller, L.M., D’Esposito, M., 2005. Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. *J. Neurosci.* 25 (25), 5884–5893. <https://doi.org/10.1523/JNEUROSCI.0896-05.2005>.
- Mitchell, R.L.C., 2006. How does the brain mediate interpretation of incongruent auditory emotions ? The neural response to prosody in the presence of conflicting lexico-semantic cues. *Eur. J. Neurosci.* 24 (12), 3611–3618. <https://doi.org/10.1111/j.1460-9568.2006.05231.x>.
- Müller, V.I., Cieslik, E.C., Kellermann, T.S., Eickhoff, S.B., 2014. Crossmodal emotional integration in major depression. *Soc. Cogn. Affect. Neurosci.* 9 (6), 839–848. <https://doi.org/10.1093/scan/nst057>.
- Müller, V.I., Habel, U., Derntl, B., Schneider, F., Zilles, K., Turetsky, B.I., Eickhoff, S.B., 2011. Incongruence effects in crossmodal emotional integration. *Neuroimage* 54 (3), 2257–2266. <https://doi.org/10.1016/j.neuroimage.2010.10.047>.
- Murty, V.P., Ritchey, M., Adcock, R.A., LaBar, K.S., 2010. fMRI studies of successful emotional memory encoding: a quantitative meta-analysis. *Neuropsychologia* 48 (12), 3459–3469. <https://doi.org/10.1016/j.neuropsychologia.2010.07.030>.
- Neufeld, J., Hagström, A., Westeinde, A.V., Lundin, K., Cauvet, E., Willfors, C., Isaksson, J., Lichtenstein, P., Bölte, S., 2020. Global and local visual processing in autism – a cotwin-control study. *JCPP (J. Child Psychol. Psychiatry)* 61 (4), 470–479. <https://doi.org/10.1111/jcpp.13120>.
- Newman-Norlund, R.D., Schie, H.T.van, Zuijlen, A.M.J.van, Bekkering, H., 2007. The mirror neuron system is more active during complementary compared with imitative action.. *Nat. Neurosci.* 10 (7), 817–818. <https://doi.org/10.1038/nn1911>.
- Nielson, D.M., Smith, T.A., Sreekumar, V., Dennis, S., Sederberg, P.B., 2015. Human hippocampus represents space and time during retrieval of real-world memories. *Proc. Natl. Acad. Sci. U. S. A.* 112 (35), 11078–11083. <https://doi.org/10.1073/pnas.1507104112>.
- Nieminen-von Wendt, T., Metsähonkala, L., Kulomäki, T., Aalto, S., Autti, T., Vanhala, R., von Wendt, L. (2003). Changes in cerebral blood flow in Asperger syndrome during theory of mind tasks presented by the auditory route. *Eur. Child Adolesc. Psychiatr.* 12(4), 178–189. <https://doi.org/10.1007/s00787-003-0337-z>.
- Ochsner, K.N., Hughes, B., Robertson, E.R., Cooper, J.C., Gabrieli, J.D., 2009. Neural systems

- supporting the control of affective and cognitive conflicts. *J. Cognit. Neurosci.* 21 (9), 1842–1855. <https://doi.org/10.1162/jocn.2009.21129>.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9 (1), 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4).
- Park, J.Y., Gu, B.M., Kang, D.H., Shin, Y.W., Choi, C.H., Lee, J.M., Kwon, J.S., 2010. Integration of cross-modal emotional information in the human brain: an fMRI study. *Cortex* 46 (2), 161–169. <https://doi.org/10.1016/j.cortex.2008.06.008>.
- Peelen, M.V., Atkinson, A.P., Vuilleumier, P., 2010. Supramodal representations of perceived emotions in the human brain. *J. Neurosci.* 30 (30), 10127–10134. <https://doi.org/10.1523/JNEUROSCI.2161-10.2010>.
- Piggot, J., Kwon, H., Mobbs, D., Blasey, C., Lotspeich, L., Menon, V., Bookheimer, S., Reiss, A.L., 2004. Emotional attribution in high-functioning individuals with autistic spectrum disorder: a functional imaging study. *J. Am. Acad. Child Adolescent Psychiatr.* 43 (4), 473–480. <https://doi.org/10.1097/00004583-200404000-00014>.
- Robertson, A.E., Simmons, D.R., 2013. The relationship between sensory sensitivity and autistic traits in the general population. *J. Autism Dev. Disord.* 43 (4), 775–784. <https://doi.org/10.1007/s10803-012-1608-7>.
- Rubia, K., Hyde, Z., Halari, R., Giampietro, V., Smith, A., 2010. Effects of age and sex on developmental neural networks of visual–spatial attention allocation. *Neuroimage* 51 (2), 817–827. <https://doi.org/10.1016/j.neuroimage.2010.02.058>.
- Rutherford, M.D., 2007. The “Reading the Mind in the Voice” test-revised: a study of complex emotion recognition in adults with and without autism spectrum conditions. *J. Autism Dev. Disord.* 37 (6), 1096–1106. <https://doi.org/10.1007/s10803-006-0252-5>.
- Samson, F., Mottron, L., Soulières, I., Zeffiro, T.A., 2012. Enhanced visual functioning in autism: an ALE meta-analysis. *Hum. Brain Mapp.* 33 (7), 1553–1581. <https://doi.org/10.1002/hbm.21307>.
- Scharpf, K.R., Wendt, J., Lotze, M., Hamm, A.O., 2010. The brain’s relevance detection network operates independently of stimulus modality. *Behav. Brain Res.* 210 (1), 16–23. <https://doi.org/10.1016/j.bbr.2010.01.038>.
- Schelinski, S., Riedel, P., von Kriegstein, K., 2014. Visual abilities are important for auditory-only speech recognition: evidence from autism spectrum disorder. *Neuropsychologia* 65, 1–11. <https://doi.org/10.1016/j.neuropsychologia.2014.09.031>.
- Schelinski, S., Roswandowitz, C., von Kriegstein, K., 2017. Voice identity processing in autism spectrum disorder. *Autism Res.* 10 (1), 155–168. <https://doi.org/10.1002/aur.1639>.

- Silk, T., Vance, A., Rinehart, N., Egan, G., Boyle, M.O., Br, J.L., 2005. Fronto-parietal activation in attention-deficit hyperactivity disorder, combined type : functional magnetic resonance imaging study. *Br. J. Psychiatry* 187 (3), 282–283. <https://doi.org/10.1192/bjp.187.3.282>.
- Singh-Curry, V., Husain, M., 2009. The functional role of the inferior parietal lobe in the dorsal and ventral stream dichotomy. *Neuropsychologia* 47 (6), 1434–1448. <https://doi.org/10.1016/j.neuropsychologia.2008.11.033>.
- Smith, D.G., Jones, P.S., Bullmore, E.T., Robbins, T.W., Ersche, K.D., 2013. Cognitive control dysfunction and abnormal frontal cortex activation in stimulant drug users and their biological siblings. *Transl. Psychiatry* 3 (5), e257. <https://doi.org/10.1038/tp.2013.32>.
- Su, W., Culotta, M., Mueller, J., Tsuzuki, D., Pelphrey, K., Id, A.B., 2020. Differences in cortical activation patterns during action observation , action execution , and interpersonal synchrony between children with or without autism spectrum disorder (ASD): an fNIRS pilot study. *PLoS One* 15 (10). <https://doi.org/10.1371/journal.pone.0240301>.
- Talairach, J., Tournoux, P., 1988. *Co-planar Stereotaxic Atlas of the Human Brain 3-Dimensional Proportional System: An Approach to Cerebral Imaging*. Thieme, New York/Stuttgart.
- Thinnes-Elker, F., Iljina, O., Apostolides, J.K., Kraemer, F., Schulze-bonhage, A., Aertsen, A., Ball, T., 2012. Intention concepts and brain-machine interfacing. *Front. Psychol.* 3. <https://doi.org/10.3389/fpsyg.2012.00455>.
- Tunik, E., Rice, N.J., Hamilton, A., Grafton, S.T., 2007. Beyond grasping: representation of action in human anterior intraparietal sulcus. *Neuroimage* 36 (Suppl. 2), T77–T86. <https://doi.org/10.1016/j.neuroimage.2007.03.026>.
- Von dem Hagen, E.A., Nummenmaa, L., Yu, R., Engell, A.D., Ewbank, M.P., Calder, A.J., 2011. Autism spectrum traits in the typical population predict structure and function in the posterior superior temporal sulcus. *Cerebr. Cortex* 21 (3), 493–500. <https://doi.org/10.1093/cercor/bhq062>.
- Wang, A.T., Dapretto, M., Hariri, A.R., Sigman, M., Bookheimer, S.Y., 2004. Neural correlates of facial affect processing in children and adolescents with autism spectrum disorder. *J. Am. Acad. Child Adolesc. Psychiatry* 43 (4), 481–490. <https://doi.org/10.1097/00004583-200404000-00015>.
- Wang, A.T., Lee, S.S., Sigman, M., Dapretto, M., 2007. Reading affect in the face and voice: neural correlates of interpreting communicative intent in children and adolescents with autism spectrum disorders. *Arch. Gen. Psychiatr.* 64 (6), 698–708. <https://doi.org/10.1001/archpsyc.64.6.698>.
- Watanabe, T., Yahata, N., Abe, O., Kuwabara, H., Inoue, H., Takano, Y., Iwashiro, N., Natsubori, T., Aoki, Y., Takao, H., Sasaki, H., Gono, W., Murakami, M., Katsura, M., Kunimatsu, A., Kawakubo, Y., Matsuzaki, H., Tsuchiya, K.J., Kato, N., Yamasue, H., 2012. Diminished medial prefrontal activity

- behind autistic social judgments of incongruent information. *PloS One* 7 (6), 1–11.
<https://doi.org/10.1371/journal.pone.0039561>.
- Watson, R., Latinus, M., Noguchi, T., Garrod, O., Crabbe, F., Belin, P., 2013. Dissociating task difficulty from incongruence in face-voice emotion integration. *Front. Hum. Neurosci.* 7 (744), 1–13.
<https://doi.org/10.3389/fnhum.2013.00744>.
- Wittfoth, M., Schröder, C., Schardt, D.M., Dengler, R., Heinze, H.J., Kotz, S.A., 2010. On emotional conflict: interference resolution of happy and angry prosody reveals valencespecific effects. *Cerebr. Cortex* 20 (2), 383–392. <https://doi.org/10.1093/cercor/bhp106>.
- Woodward, T.S., Ruff, C.C., Ngan, E.T.C., 2006. Short- and long-term changes in anterior cingulate activation during resolution of task-set competition. *Brain Res.* 1068 (1), 161–169.
<https://doi.org/10.1016/j.brainres.2005.10.094>.
- Woodward, T.S., Metzak, P.D., Meier, B., Holroyd, C.B., 2008. Anterior cingulate cortex signals the requirement to break inertia when switching tasks: a study of the bivalency effect. *Neuroimage* 40 (3), 1311–1318. <https://doi.org/10.1016/j.neuroimage.2007.12.049>.
- Wright, T.M., Pelphrey, K.A., Allison, T., McKeown, M.J., McCarthy, G., 2003. Polysensory interactions along lateral temporal regions evoked by audiovisual speech. *Cerebr. Cortex* 13 (10), 1034–1043.
<https://doi.org/10.1093/cercor/13.10.1034>.
- Zhang, X., Yin, L., Cohn, J.F., Canavan, S., Reale, M., Horowitz, A., Liu, P., Girard, J.M., 2014. BP4D-Spontaneous: a high-resolution spontaneous 3D dynamic facial expression database. *Image Vis Comput.* 32 (10), 692–706. <https://doi.org/10.1016/j.imavis.2014.06.002>.