Facial Affect Recognition Training in Autism: Can We Animate the Fusiform Gyrus?

Sven Bölte
J. W. Goethe University

Daniela Hubl
University Hospital of Clinical Psychiatry

Sabine Feineis-Matthews and David Prvulovic
J. W. Goethe University

Thomas Dierks
University Hospital of Clinical Psychiatry

Fritz Poustka
J. W. Goethe University

One of the most consistent findings in the neuroscience of autism is hypoactivation of the fusiform gyrus (FG) during face processing. In this study the authors examined whether successful facial affect recognition training is associated with an increased activation of the FG in autism. The effect of a computer-based program to teach facial affect identification was examined in 10 individuals with high-functioning autism. Blood oxygenation level-dependent (BOLD) functional magnetic resonance imaging (fMRI) changes in the FG and other regions of interest, as well as behavioral facial affect recognition measures, were assessed pre- and posttraining. No significant activation changes in the FG were observed. Trained participants showed behavioral improvements, which were accompanied by higher BOLD fMRI signals in the superior parietal lobule and maintained activation in the right medial occipital gyrus.

Keywords: autism, fMRI, emotion recognition, face processing, neuropsychotherapy

As soon as after some minutes of life, an innate preference for facelike visual stimuli can be observed in healthy newborns (Valenza, Simion, Macchi Cassia, & Umilta, 1996). For the communication of affect and conveyance of reciprocal social interaction, the human face is a pivotal player. A lack of facial affect recognition has been identified in a range of mental disorders (e.g., Kee, Green, Mintz, & Brekke, 2003; Kucharska-Pietura, Nikolau, Masiak, & Treasure, 2004). Numerous studies have provided evidence that limitations in the capacity to judge facially expressed emotions are likely to be a consistent part of autism (e.g., Bölte & Poustka, 2003; Bormann-Kischkel, Vilsmeier, & Baude, 1995; Hobson, 1986). The same is true for individuals with higher functioning autism and Asperger’s syndrome. However, deficits are milder and may appear rather insignificant in the first place, owing to compensatory strategies (e.g., Capps, Yirmiya, & Sigman, 1992; Grossman, Klin, Carter, & Volkmar, 2000; Teunisse & de Gelder, 2001). Despite one recent negative finding (Hadjikhani et al., 2004), imaging studies have particularly and recurrently identified decreased activation of the fusiform gyrus (FG) during face and facial-affect processing in autism spectrum disorders (e.g., Critchley et al., 2000; Hubl et al., 2003; Pierce, Muller, Ambrose, Allen, & Courchesne, 2001; Schultz et al., 2000; Wang, Dapretto, Harrir, Sigman, & Bookheimer, 2004). Additional evidence from imaging studies in autism suggests that hypoactivation of the FG could be a key cerebral correlate of an even considerably broader array of maladapted social and cognitive mechanisms in the disorder (Schultz et al., 2003).

Several intervention programs have been developed to improve facial affect understanding in autism (e.g., Silver & Oakes, 2001; Tanaka, Lincoln, & Hegg, 2003). Computers seem to be a preferred learning medium for people with autism (Chen & Bernard-Opitz, 1993; Williams, Wright, Callaghan, & Coughlan, 2002), and computer-based interventions generally appear to be promising in the disorder (Bölte, 2004). In a previous report (Bölte et al., 2002), we described the development and successful pilot evaluation of a computer-based program to teach and test the ability to identify basic facially expressed emotions on two different levels, called the Frankfurt Test and Training of Facial Affect Recognition (FEFA). Neurobiological changes associated with cognitive training or psychotherapy are viewed as one of the most topical issues in mental health research (Gabbard, 2000). Measurable effects on the brain of psychological treatment have been demonstrated for several psychiatric disorders, for instance, schizophrenia, phobia, and depression (Goldapple et al., 2004; Paquette et al., 2003; Wykes et al., 2002). To our knowledge, no study of this kind...
has yet been undertaken in autism. The purpose of this study was to probe by functional magnetic resonance imaging (fMRI) whether observed improvements regarding facial affect recognition in autism following FEFA training are accompanied by an increased activation of the FG in the brain.

Method

Participants

The investigation was approved by the local ethics committee of the University of Frankfurt, and all probands and their parents or caregivers had given written informed consent to participate in the procedures. Ten, mostly higher functioning Caucasian male adolescent and adult individuals with idiopathic autism were included in the study. They were inpatients or outpatients of the Department of Child and Adolescent Psychiatry at Frankfurt University or had been recruited within an ongoing international project on the molecular genetics of autism (www.well.ox.ac.uk/~maestrin/iat.html). None of them received any psychoactive medication during the study. All fulfilled the research criteria of the International Classification of Diseases (10th rev.; World Health Organization, 1992) as well as the diagnostic algorithm thresholds for the disorder in the Autism Diagnostic Interview—Revised (Lord, Rutter, & Le Couteur, 1994) and the Autism Diagnostic Observation Schedule Module 3 or 4 (Lord, Rutter, DiLavore, & Risi, 2001). Five of the participants were randomly assigned to the experimental group for receiving FEFA emotion recognition training; the other 5 served as a control group. The mean age in the experimental group was 29.4 (SD = 5.9) years, and their mean nonverbal IQ on the Raven Standard Progressive Matrices (Raven, 1996) was 94.3 (SD = 18.9). In the control group the average age was 25.8 (SD = 8.0) years, and average nonverbal IQ was 98.6 (SD = 19.2).

Design

The current study is a pre–post within-subject design with one experimental group (autism participants receiving facial affect recognition training) and one control group (autism participants without training). Shortly before and after the experimental group had gone through the training phase, brain images were assessed in both samples applying an fMRI activation paradigm, comprising visual stimuli being similar but not identical to the material of the FEFA. Additionally, parallel to the scanning, repeated measures on the FEFA test modules were collected. Both the FEFA test modules and imaging procedures in general have shown sufficient retest reliability to be used for repeated measuring (Bölte et al., 2002; Yetkin, McAuliffe, Cox, & Haughton, 1996). During the training phase, participants in the experimental group received FEFA training assisted by an experienced clinician over a period of 5 weeks, consisting of 2 hr training a week at the Department of Child and Adolescent Psychiatry at Frankfurt University. Despite ongoing inpatient treatment, the members of the control group did not get any comparable specific emotion-recognition treatment during the study, but they were offered FEFA training subsequently.

Materials and Procedure

FEFA. The FEFA is a computer-based program to test and teach the recognition of facially expressed basic emotions, described elsewhere in detail (Bölte et al., 2002). It uses the cross-cultural concept of seven fundamental affective states (happy, sad, angry, surprised, disgusted, fearful, and neutral), devised by Paul Ekman and colleagues (Ekman, Friesen, & Ellsworth, 1972), for judging affect in photographs of whole faces and eye regions, respectively.

The FEFA training module comprises about 500 facial affect teaching items on three description levels. On Level 1, the test and the training module use comparable stimulus material, but correct judgments in the training section are followed by visual and acoustical reinforcement. If the given answer is incorrect, a feedback button appears on the screen. Clicking on the link leads to an explanation of the item solution (Level 2). A further in-depth engagement in the specific emotion is provided by the opportunity to look at a comic strip and again choose the specific corresponding emotion (Level 3).

The normed FEFA test module comprises a series of 50 items for faces and 40 items for eyes, showing good to excellent inter-rater and test-retest reliability (r < .89). As 1 point is given for each correctly classified emotion, the maximum score is 50 for the face test and 40 for the eyes test. Mean normative values on the FEFA test are 42.9 (SD = 1.9) for the face test and 34.5 (SD = 2.2) for the eyes test.

MRI procedure. Magnetic resonance images were acquired using a 1.5-tesla whole-body Siemens Magnetom Vision system (Siemens Medical Systems, Erlangen, Germany) at the Department of Neuroradiology at Frankfurt University. A high-resolution, T1-weighted three-dimensional whole-brain anatomic data set (MP RAGE) was gathered for each individual (voxel size 1 × 1 × 1 mm). For functional imaging, 128 volumes were acquired (1 volume = 15 axial slices covering the whole brain; repetition time/echo time = 4,000 ms/69 ms; flip angle = 90°; field of view = 210 × 210 mm; slice thickness = 5 mm; interslice distance = 1 mm; voxel size = 1.6 × 1.6 × 5 mm). The fMRI activation paradigm followed a classical block design. Each trial was built from eight activation blocks alternating with eight resting blocks, serving as baseline. Activation blocks were presented in a pseudorandom order, each consisting of eight scanning volumes lasting 32 s. Standardized photographs of faces exhibiting basic emotional states (happy, sad, angry, neutral; Ekman & Friesen, 1979) as well as generated scrambled versions of them were displayed in the scanner. Blocks of emotional faces were shown connected with two different tasks: pressing a button with the right index finger when a happy face appeared (three blocks) or when a face of a woman appeared (three blocks). Each task was indicated by using a one-word cue at the beginning of each block. To control for attention differences, one of the scrambled task blocks had the same instruction as the gender task and the other had the same instruction as the emotion task.

Data Analysis

fMRI data processing. For magnetic resonance data analysis, registration, and visualization, the fMRI software package BrainVoyager 2000 (BrainInnovation, Maastricht, the Netherlands) was used. After temporal and spatial smoothing using a default parameter, the two-dimensional statistical maps were superimposed onto the original functional scans and incorporated into the three-dimensional anatomical data sets through interpolation of the functional voxels to the same resolution as the anatomical voxels (Dierks et al., 1999). Aside from the FG (Brodmann area [BA] 37), three additional regions of interest (ROIs) in both hemispheres were defined according to the results gained in our previous study on face processing in autism (Hubl et al., 2003): the medial occipital gyrus (GOm, BA 19), the superior parietal lobule (SPL, BA 19), and the precentral gyrus (GPreC, BA 6).

In a first step we tested whether all ROIs identified in our previous study were relocated in the present study, too. Statistical analysis of the variance of the mean blood oxygenation level-dependent (BOLD) signal was based on the application of multiple regression analysis to time series of task-related functional activation (Friston et al., 1995). A general linear model of the experiment was computed from the seven (one for each participant) tagged and z-normalized volume time courses. The signal values during the face tasks were considered the effects of interest. The corresponding predictors, obtained by convolution of an ideal boxcar response (assuming the value 1 for the time points of task presentation and the value 0 for the remaining time points) with a linear model of the hemodynamic response (Boonyton, Engel, Glover, & Heeger, 1996), were used to build the design.
matrix of the experiment. The global level of the signal time courses in each session was considered to be a confounding effect.

In a second step, to analyze the effects for the activation paradigm tasks compared with baseline, three-dimensional statistical maps were generated for each subject analysis on the basis of individually computed correlations of the BOLD time course, with the time course of the block design paradigm serving as predictor. In the RoIs, voxels were accepted as activated only when the associated p value was < .01 (corresponding to a correlation coefficient r > .35) and when they formed part of a cluster of 200 mm³ or more. The mean signal time course was used for further analysis.

Analysis of behavioral and magnetic resonance data. Trained and untrained individuals with autism were compared pre- and posttraining on two behavioral measures (the FEFA face and eyes tests) and BOLD signal changes in four RoIs (FG, GOm, SPL, GPreC). Separate analyses were conducted for the repeated measures on the behavioral and the imaging level. All inferential statistics were computed using SPSS 11.5 for Windows. Simultaneous analyses of covariance (ANCOVAs) for repeated measures were calculated to determine the training effect. Herein, group (trained vs. untrained) was a between-subjects factor, and age and IQ were inserted as covariates. On the behavioral level, performance on the FEFA face and eyes tests served as a within-subject factor. On the imaging level, percentage BOLD signal changes in each right and left hemisphere RoI were compared within the individuals. Because of the small sample size and the increased risk for Type II errors, for all statistics alpha level was set at p < .10 and large effect sizes of \( \eta^2 > .50 \) were also interpreted as significant. In addition to the group comparisons, the correlations between BOLD signal changes and behavioral changes on the FEFA face and eyes tests were computed in the experimental group using Spearman’s coefficient.

Results

Functional Imaging Data

Figure 1 shows the resulting contrast maps (stimulation vs. rest) for the general linear model for all participants and conditions. Groups’ results are summarized in Table 1. Owing to noncorrectable movement artifacts during fMRI, complete scanning data were not available for all probands for all scans carried out. Unfortunately, 2 participants from the experimental group and 1 from the control group had to be excluded from the imaging data analysis. In the remaining participants, no significant BOLD signal changes were found in the FG pre- and posttraining. On the other hand, ANCOVAs revealed significant group differences for two other RoIs, namely GOm right, \( F(1, 3) = 16.1, p = .03, \eta^2 = .84 \), and SPL right, \( F(1, 3) = 4.1, p = .18, \eta^2 = .67 \). The effect regarding GOm right was almost exclusively due to a decreased activation in the control group at posttreatment scan, whereas it was primarily due to an increased activation in the experimental group in SPL right. Hemodynamic response functions (faces vs. rest) pre- and posttraining for the trained and untrained group for all RoIs are shown in Figure 2.

Behavioral Data

Mean score at baseline in the experimental group was 31.6 (SD = 9.9) for the face test and 17.6 (SD = 6.7) for the eyes test. At follow-up the values were 43.0 (SD = 3.2) for the face test and 31.2 (SD = 4.9) for the eyes test, an improvement of more than one standard deviation for the face test and more than two for the eyes test. Each proband in the trained group showed an increase in affect recognition from pre- to posttesting on both measures. The correlations between the behavioral improvements on the FEFA face and eyes tests and the BOLD signal changes were .67 (eyes) and .56 (face) for GOm right and .75 (eyes) and .76 (face) for SPL right.

In the control group, the results were almost identical for the first and the second measures. Here, the mean score was 33.6 (SD = 9.9) for the face test and 21.4 (SD = 9.8) for the eyes test at baseline and 33.0 (SD = 9.4) for the face test and 21.4 (SD = 9.5) for the eyes test at follow-up. ANCOVA routines for repeated measures yielded significantly enhanced performances in the

![Figure 1](image-url)
Figure 2. Hemodynamic response functions (faces vs. rest) for regions of interest pre- and posttraining in the trained (dark lines) and untrained (light lines) groups. Dashed line signifies pretraining; solid line signifies posttraining. BOLD = blood oxygenation level-dependent; FG = fusiform gyrus; GOm = medial occipital gyrus; GPreC = precentral gyrus; SPL = superior parietal lobule; r = rest; s = stimulus.

Table 1
Pre- and Posttraining Means and Standard Deviations for BOLD Signal Changes (%) in Trained and Untrained Individuals With Autism and ANCOVA Inference Statistics Results for Group Comparisons

<table>
<thead>
<tr>
<th>Region</th>
<th>M (SD)</th>
<th>ANCOVA</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trained group</td>
<td>Untrained</td>
<td>F</td>
<td>p</td>
<td>η²</td>
</tr>
<tr>
<td>FG right pre</td>
<td>1.38 (0.59)</td>
<td>1.13 (0.32)</td>
<td>0.38</td>
<td>.58</td>
<td>.11</td>
</tr>
<tr>
<td>FG right post</td>
<td>1.04 (0.26)</td>
<td>1.07 (0.16)</td>
<td>1.8</td>
<td>.27</td>
<td>.38</td>
</tr>
<tr>
<td>FG left pre</td>
<td>1.14 (0.20)</td>
<td>1.07 (0.21)</td>
<td>1.42 (0.59)</td>
<td>0.95 (0.29)</td>
<td>16.1</td>
</tr>
<tr>
<td>FG left post</td>
<td>1.08 (0.49)</td>
<td>1.11 (0.31)</td>
<td>0.38</td>
<td>.58</td>
<td>.11</td>
</tr>
<tr>
<td>GOm right pre</td>
<td>1.41 (0.36)</td>
<td>1.46 (0.48)</td>
<td>1.64 (0.62)</td>
<td>1.58 (0.42)</td>
<td>2.4</td>
</tr>
<tr>
<td>GOm right post</td>
<td>1.42 (0.59)</td>
<td>0.95 (0.29)</td>
<td>16.1</td>
<td>.03*</td>
<td>.84*</td>
</tr>
<tr>
<td>GOm left pre</td>
<td>1.64 (0.62)</td>
<td>1.58 (0.42)</td>
<td>2.4</td>
<td>.22</td>
<td>.45</td>
</tr>
<tr>
<td>GOm left post</td>
<td>1.66 (0.61)</td>
<td>1.01 (0.36)</td>
<td>16.1</td>
<td>.03*</td>
<td>.84*</td>
</tr>
<tr>
<td>SPL right pre</td>
<td>0.58 (0.16)</td>
<td>0.86 (0.45)</td>
<td>1.6</td>
<td>.30</td>
<td>.35</td>
</tr>
<tr>
<td>SPL right post</td>
<td>1.10 (0.38)</td>
<td>0.72 (0.28)</td>
<td>0.99 (0.43)</td>
<td>0.86 (0.13)</td>
<td>0.03</td>
</tr>
<tr>
<td>SPL left pre</td>
<td>0.95 (0.16)</td>
<td>1.13 (0.42)</td>
<td>1.28 (0.30)</td>
<td>0.89 (0.28)</td>
<td>0.39</td>
</tr>
<tr>
<td>SPL left post</td>
<td>1.10 (0.38)</td>
<td>0.72 (0.28)</td>
<td>0.99 (0.43)</td>
<td>0.86 (0.13)</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Note. BOLD = blood oxygenation level-dependent; ANCOVA = analysis of covariance; FG = fusiform gyrus; GOm = medial occipital gyrus; SPL = superior parietal lobule; GPreC = precentral gyrus.

*a n = 3. b n = 4.*

* Considered significant differences (p < .10 or η² > .50).
trained compared with the untrained autism sample for both the face test, $F(1, 16) = 8.5, p = .013, \eta^2 = .59$, and the eyes test, $F(1, 16) = 42.9, p = .0005, \eta^2 = .88$.

Discussion

Previous studies have consistently reported a hypoactivation of the FG during face processing in autism. We hypothesized an increased activation in the FG following successful facial affect recognition training on a behavioral level. In line with other findings (e.g., Silver & Oakes, 2001), we found individuals with high-functioning autism to profit from affect recognition training. In this study, intensively trained individuals with autism exhibited considerable improvements in basic emotion detection skills and mostly reached normative values on the FEFA eyes and face tests after the FEFA intervention. Nevertheless, contrary to our expectation, no associated activation changes were found in the FG. Improved facial affect recognition performance was accompanied by higher activation of the right SPL. Moreover, whereas untrained participants exhibited a marked decrease in GOM right activation at second scan, trained participants showed maintained alertness in this area. Both SPL right and GOM right have been assumed to be part of a compensatory facial processing network (Hubl et al., 2003). Whereas the GOM has generally been described as an important region for object and face recognition (e.g., Malach et al., 1995), SPL is considered to be especially involved in visuospatial skills (e.g., Sack et al., 2002) as well as visual attention (Wojciulik & Kanwisher, 1999). Thus, against our hypothesis, the fMRI results indicate that the observed improvements in facial affect recognition on the behavioral level correlate with compensatory mechanisms, not with an animation of the FG.

The present study has several limitations that perhaps compromise the validity of our data interpretation. The most important is the small sample size, particularly given that the main prediction was not supported. There is a definite need for replication of our fMRI findings in a larger sample, also including a normative control group. Moreover, although we found significant improvements on the behavioral level, we have no hard empirical data on the clinical significance and generalization of these effects. Although our clinical experience and parental feedback indicate at least some sort of transfer into daily life, it cannot be ruled out that the effects reflect familiarity with the task or general behavior changes in attention and memory, not true gains in facial affect understanding. The latter possibility has been suggested by other authors (e.g., Grossman et al., 2000; Teunisse & de Gelder, 2001). Generalization is also limited in terms of age, as our sample did not include children, only adolescents and adults. Older individuals might be less open to treatment, and younger individuals might be more likely to adopt typical mechanisms, including FG activation. Finally, a recent study by Dalton et al. (2005) suggests that diminished gaze fixation may account for fusiform hypopactivation in autism. Unfortunately, we did not control for gaze fixation using an eye tracking device in this study.

In summary, the present study is the first to explore the neurobiological correlates of neuropsychological intervention in autism. We found that gains in facial affect recognition in autism are associated with higher activation in brain areas being part of a compensatory facial processing network, not necessarily with activation changes in the FG. However, because of the small sample size and other limitations, our study should probably be looked at as heuristic and generating hypotheses in this research field rather than stringently testing them.

References


Received April 13, 2005
Revision received September 13, 2005
Accepted October 28, 2005