OXFORD

Social Cognitive and Affective Neuroscience, 2017, 1711–1719

doi: 10.1093/scan/nsx117
Advance Access Publication Date: 4 October 2017
Original article

Anger-sensitive networks: characterizing neural systems recruited during aggressive social interactions using data-driven analysis

Frederike Beyer,[1,2,3] Ulrike M. Krämer,[1,2] and Christian F. Beckmann[4,5,6]

[1]Department of Neurology, [2]Institute of Psychology II, University of Lübeck, 23538 Lübeck, Germany, [3]Institute of Cognitive Neuroscience, University College London, London, UK, [4]Department of Cognitive Neuroscience, Radboud University Medical Centre, Nijmegen, The Netherlands, [5]Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands, and [6]Nuffield Department of Clinical Neurosciences, Oxford Centre for Functional MRI of the Brain, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DU, UK

Ulrike M. Krämer and Christian F. Beckmann authors contributed equally to this work.
Correspondence should be addressed to Frederike Beyer, Institute of Cognitive Neuroscience, University College London, Alexandra House, 17 Queen Square, London WC1N 3AR, UK. E-mail: f.beyer@ucl.ac.uk.

Abstract

Social neuroscience uses increasingly complex paradigms to improve ecological validity, as investigating aggressive interactions with functional magnetic resonance imaging (fMRI). Standard analyses for fMRI data typically use general linear models (GLM), which require a priori models of task effects on neural processes. These may inadequately model non-stimulus-locked or temporally overlapping cognitive processes, as mentalizing about other agents. We used the data-driven approach of independent component analysis (ICA) to investigate neural processes involved in a competitive interaction. Participants were confronted with an angry-looking opponent while having to anticipate the trial outcome and the opponent's behaviour. We show that several spatially distinctive neural networks with associated temporal dynamics were modulated by the opponent's facial expression. These results dovetail and extend the main effects observed in the GLM analysis of the same data. Additionally, the ICA approach identified effects of the experimental condition on neural systems during inter-trial intervals. We demonstrate that cognitive processes during aggressive interactions are poorly modelled by simple stimulus onset/duration variables and instead have more complex temporal dynamics. This highlights the utility of using data-driven analyses to elucidate the distinct cognitive processes recruited during complex social paradigms.

Key words: aggression; fMRI; independent component analysis

Introduction

The field of social neuroscience is a research area undergoing rapid changes. Recent emphasis is placed on external validity of experimental designs that aim at investigating social interactive processes (Schilbach *et al.*, 2013). Consequently, more complex experimental designs are starting to replace standardized stimulus presentation and fixed stimulus-response designs.

In previous research, we have used such designs to shed light on the neural basis of aggression in healthy adults. Specifically, we have adapted the well-established Taylor Aggression Paradigm (TAP; Taylor, 1967) for use with functional

Received: 14 June 2017; **Revised:** 9 September 2017; **Accepted:** 1 October 2017

© The Author (2017). Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

magnetic resonance imaging (fMRI; Krämer *et al.*, 2007) and electroencephalography (Krämer *et al.*, 2008). The TAP is set up as a competitive reaction time task, in which the participant competes with an ostensible opponent. In each round, the loser gets punished with an aversive stimulus, the intensity of which is determined by the winner. Using this approach, we studied, for instance, the effects of psychopharmacological manipulations on aggression and its neural underpinnings (Krämer *et al.*, 2011), influences of inter-individual variability in sensitivity to threat on mentalizing networks (Beyer *et al.*, 2014) and effects of trait aggressiveness on decision-related frontal activity during aggressive encounters (Krämer *et al.*, 2008; Krämer *et al.*, 2009). Further and in order to investigate the role of the neural processing of angry facial expressions during aggressive interactions, we implemented video feedback of an opponent in the TAP. Data analyses of this study within the framework of classical general linear modelling (GLM) established a negative correlation between orbitofrontal activity to angry *vs* neutral faces and participants' aggressive behaviour (Beyer *et al.*, 2015).

However, traditional data analysis strategies may be inadequate for the use with such complex designs. The widely used GLM approach requires the definition of onset regressors for stimulus presentations and responses. In simple stimulus-response designs, this approach is assumed to model most of the underlying task-related processes. Despite its wide use and popularity, the GLM-approach can be problematic, for example if the chosen number of regressors is too small or too large to optimally model the task-related variance of the signal (Poline and Brett, 2012). In social interactions, many processes can be assumed to occur throughout the task, without being strictly stimulus or response-locked. Within the GLM framework, BOLD-signal changes related to such processes may be unobservable, due to the lack of an a priori model of the process-related timings.

In the case of the TAP, such simultaneous cognitive processes involve the evaluation of opponent-related information, selecting a punishment level and anticipating the following task stages. Given the difficulty in pinpointing these processes to a specific onset, analysis approaches using fixed models of stimulus-driven BOLD-changes may fail to characterize the neural processes actually underlying behaviour in the TAP.

Further, in simple experimental designs, temporal derivatives may be used to account for shifts between task onset and BOLD-response. In tasks as the TAP, where many cognitive processes likely occur in short succession and may partially overlap, adding temporal derivatives to a large number of regressors will result in design collinearity. The shortfalls of GLM-based analysis for disentangling different task contributions is largely due to the univariate nature of the approach, where individual voxels are being tested against an experimental design. Multi-variate approaches, by comparison, are able to incorporate the information across multiple voxels' time series and therefore will exploit spatially distributed statistical regularities in the data. Investigating the entire time-course of the TAP in this manner should be helpful to reveal effects that do not fit the a priori model.

Independent component analysis (ICA) constitutes such an approach. This method identifies spatial patterns of BOLD-signal changes across the entire brain and the entire experiment time-course (McKeown *et al.*, 1998) based on the assumption of statistical independence between these maps. As such, this method does not implement any a-priori assumptions as to onsets or locations of task effects. Due to this model-free nature, ICA is especially useful for artefact detection, denoizing of data and analysing resting state data (Beckmann, 2012). For the same reason, it is also useful for the analysis of task-related data that consist of long and inhomogeneous task periods (Beckmann, 2012), as for example, during motor sequence learning (Kincses *et al.*, 2008). ICA can be used to explore spatial network properties as the composition of the default mode network (Leech *et al.*, 2011) and to explore temporal dynamics of networks across task onset and offset (Scott *et al.*, 2015).

In this study, we explored the use of probabilistic ICA (Beckmann and Smith, 2004) on complex, task-fMRI data derived from our version of the TAP where the opponent showed neutral and angry facial expressions. We re-analysed data previously studied with a GLM to test the validity of the approach by directly comparing ICA and GLM results and to verify additional information gained by the model-free analysis. In a first step, we identified task-related networks among all ICA components to identify those that are recruited during the aggressive interaction. We further investigated how these relate to different task conditions, i.e. whether activity in these networks was modulated by angry *vs* neutral facial expressions of the opponent. Finally, the GLM analyses revealed a correlation between activity in the dorsal anterior cingulate cortex (dACC) and intra-subject variability in aggressive behaviour in angry trials (Beyer *et al.*, 2015). Thus, we also related component activity to between-trial, within-subject variability in aggressive behaviour, to show whether the recruitment of a neural system including the dACC would be differentially modulated during trials in which participants chose to aggress or refrained from aggression.

Thus, the aim of this study was to allow for a direct comparison between GLM and data-driven analysis approaches. We hypothesized that the ICA should uncover networks reflecting activation patterns observed in the GLM analysis. That is, one or more networks spanning mPFC and superior temporal gyrus should show increased activity in early task stages of angry trials. Activity patterns observed for the outcome phase (increased activity in angry trials across areas in temporal parietal and prefrontal cortex) should likely split up into subnetworks. A network including dACC should be sensitive to punishment selections on a per-trial basis.

Materials and methods

Participants and experimental procedure

Forty-one male healthy volunteers participated in the study. Upon arrival in the laboratory, participants received instructions together with their ostensible opponent. They then entered the scanner for the aggression task (see below). After scanning, participants completed a post-experimental questionnaire assessing potential suspicion concerning their opponent's involvement in the task. Afterwards, participants were fully debriefed and paid for their participation.

One participant was excluded from data analysis due to excessive movement (>4 mm) during functional scanning and two due to pathological findings in their anatomical images. Five participants were excluded because they reported suspicion about their opponent's participation in the TAP and one participant was excluded, because he selected a punishment level in less than one-third of trials. Thus, 32 participants (mean age = 23.3 years, ±2.7 years) were included in the GLM analyses. For the analysis presented here, one additional participant was excluded due to poor whole-brain coverage in the functional

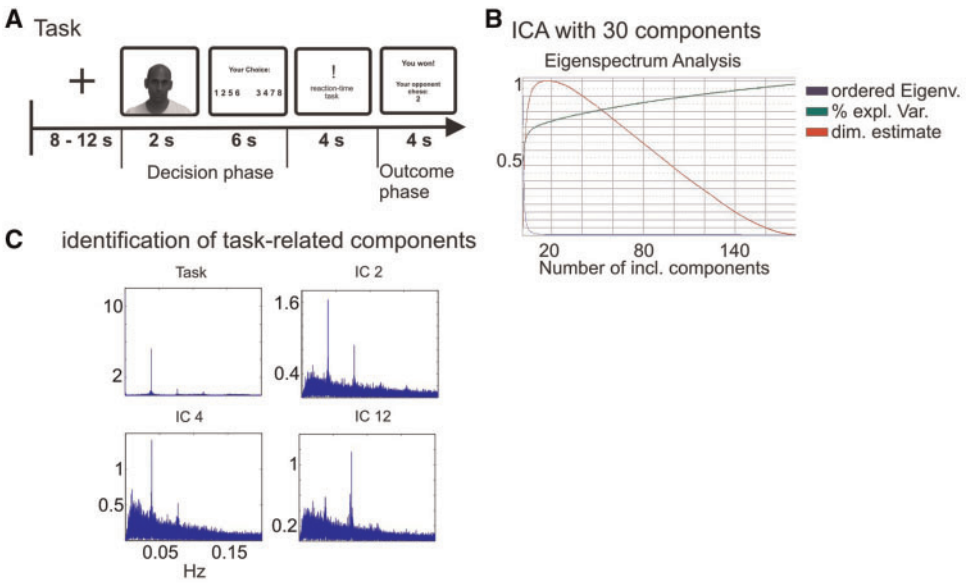

Fig. 1. Methods. The outline of a single trial of the TAP is shown in (A). (B) The eigenspektrum analysis for the estimation of the optimal number of components. (C) The frequency analyses for the task time-course and example component time-courses are shown. (D) The task regressors for the early and late half of the task, convolved with an HRF, overlaid on the boxcar function for the entire trial.

MRI data.[1] Thus, data of 31 subjects were included in this analysis.

All participants gave written informed consent and received 8 Euro per hour as compensation for their participation. The study was approved by the Ethics Committee of the University of Lübeck and performed according to the Declaration of Helsinki.

Aggression task

A detailed description of the task setup can be found in Beyer *et al.* (2015). Briefly, participants were introduced to a confederate of the experimenter who they believed would be playing against them in a competitive reaction time task, the TAP. We used a version of the TAP adapted for the use in fMRI-research. The game is split up into a decision phase in which the participant selects a punishment level for his opponent, the reaction time task, and an outcome phase in which participants learn whether he won or lost as well as learn the punishment level selected by the opponent. If the participant loses, he receives the punishment stimulus at the end of the outcome phase. We used an aversive noise as punishment, which could be adjusted in terms of loudness. This noise was adjusted individually such that, with the scanner running a functional scan, participants could clearly perceive the different noise levels, and judged the loudest noise as uncomfortable, but non-painful. At the beginning of each decision phase, we implemented video sequences showing the opponent during the punishment selection, bearing either a neutral or angry facial expression. An outline of a single trial for the TAP is given in Figure 1A. Two-second video sequences of the opponent were presented at the beginning of each trial, before participants made their punishment selection. In one-third of trials, the opponent showed an angry expression while making the punishment selection. In two-third of trials, he showed a neutral expression.

1 For multi-variate analysis of fMRI data, whole-brain coverage is of greater importance than for GLM analyses, in which independent tests are performed for each voxel.

MRI data acquisition

Structural and functional MRI images were recorded on a Philips Achieva 3-T scanner (Philips Healthcare, the Netherlands) with a standard 8-channel head coil. Functional images (624 volumes) were acquired in ascending order using a single-shot gradient-echo echo-planar imaging (EPI) sequence sensitive to blood oxygen level dependent (BOLD) contrast (TR = 2500 ms, TE = 25 ms, flip angle 80°, in-plane resolution 2.5×2.5 mm^2, image matrix 80×80, FOV 200 mm, slice thickness 2.5 mm, 47 transversal slices, SENSE factor 2.0). High resolution structural images were obtained by means of a T1-weighted 3D turbo gradient echo sequence with SENSE acceleration (SENSE factor 2; FOV = 240 mm; matrix = 240×240; 180 sagittal slices of 1 mm thickness).

Data pre-processing

Data pre-processing was performed using the FSL5.0 toolbox FEAT (Jenkinson *et al.*, 2012). Pre-processing included the following steps: brain extraction; motion correction; spatial smoothing with a 5 mm kernel; registration to a brain-extracted structural image and a standardized image in MNI-space.

To further remove movement artefacts from the functional data, we used the toolbox for Automatic Removal Of Movement Artefacts (AROMA; Pruim *et al.*, 2015). This toolbox uses FSL's MELODIC (Beckmann and Smith, 2004) toolbox to run ICA on single-session data. Movement-related components are then automatically identified and removed from the data by means of spatial regression. The cleaned data were then high-pass filtered with a 100-s frequency cut-off using FEAT.

Independent component analysis

The main steps of the analysis procedure are shown in Figure 1. The pre-processed functional images were submitted to an ICA using the FSL5.0 toolbox MELODIC. To ensure comparability of components between subjects, we used the concatenation option to conduct a group-level ICA. Thus, all functional

datasets were concatenated in time and a single ICA decomposition was generated for the entire dataset (Calhoun *et al.*, 2001). After inspection of the initial eigenspectrum analysis performed by MELODIC (Figure 1B), we fixed the number of components to be extracted to 30, a value at the upper limit of the optimal range estimation. Spatial maps were oriented according to their distribution of z-scores such that the maximum absolute z-scores had positive values, and were thresholded at $z > 3.0$. This method can result in task-negative component time courses, reflecting a relative deactivation of voxels weighted positively for respective component. Of these 30 components, task-related, non-artefact components were identified in the following steps: we visually inspected all components and excluded those representing common patterns of motion, physiological artefacts and high frequency noise artefacts (Beckmann, 2012; $n = 12$). For the remaining components ($n = 18$), we imported the component time courses, as given by the melodic toolbox, into MATLAB. Thus, for each component, we obtained a single time course, with all trials from all participants concatenated. Each data point in these time courses corresponds to one MRI volume. Using MATLAB's fast Fourier transform (fft) function, we performed a Fourier transform of the component time courses to identify the characteristic frequency of each component's time-course. For each component, we determined the frequency of maximal power.

We also computed a Fourier transform for the task time course, with all trials across all participants concatenated, which, given the timing of the task (trial duration of 24–28 s) resulted in a peak in frequency power at 0.04 Hz (Figure 1C). We then compared the peak frequency of the task time course to the peak frequencies of the component time courses.

We used this approach of spectral matching to explore, for each component, whether its signal time course was related to the task (i.e. oscillating in the same frequency as the task), without explicit modelling of task onsets. Thus, a component time course could be identified as task-related, even if it was not phase-locked to a specific event in the task or the task onset itself. Sixteen components showed a maximum of frequency power at 0.04 Hz and were thus included in further analysis as task related (Figure 1C). The remaining two components (components 12 and 13) showed a maximum of frequency power at 0.08 Hz. As shown in Figure 3, the signal time courses of these components showed two large peaks within a trial, during early and late task stages, and we therefore included them in the data analysis as well.

Time course analysis

For the 18 task-related components, we conducted time course analyses to test:

a. whether a component was differentially modulated during angry and neutral trials;
b. whether activation of a component was related to between-trial variability in punishment selection.

We did not perform a dual regression analysis, but used the original component time courses, which are based on a concatenation of the subject-wise data, to extract subject specific time courses. The concatenated signal time course for each component was split up into subject-specific sections, thus resulting in 31 time-courses per spatial group component. Each subject-wise time course consisted of 624 data points, corresponding to three runs of 208 MRI volumes.

For the visualization of component time courses, we set up two boxcar functions for the task by splitting each trial into an early phase (volumes 1–3 after task onset) and a late phase (volumes 4–6 after task onset). These boxcar functions were convolved with a hemodynamic response function and served as reference for the visualization of component time courses.

For each subject, time courses were split up into trial-wise time courses of 15 volumes each (corresponding to 37.5 s following trial onset). Note that volumes 13–15, which overlapped with the beginning of the next trial, were included for visualization only and were not included in further analysis.

To test which component time courses were differentially modulated by angry and neutral trials, we conducted paired *t*-tests for time points 3–12 following trial onset. Time points at the beginning of each trial, prior to any plausibly expected hemodynamic modulation (1 and 2), and later time points overlapping with following trials (13, 14 and 15) were omitted from this analysis. For each time point, we computed the subject-wise means for angry and neutral trials. We then compared angry against neutral trials across subjects. For statistical analyses, we used a threshold of $P < 0.05$, Bonferroni-corrected for the number of components and time points tested [$n = 180$ (18 components of 10 time points each); corrected $P < 0.0003$].

As the GLM-analysis showed a correlation of neural reactivity to angry facial expressions and aggressive behaviour within subjects, we additionally tested whether this finding could be replicated using the component time courses. Note that as this analysis is based on previous findings in the same datasets, significant results should not be understood as a replication of previous findings. Rather, this analysis serves better understanding of potential differences in the results gained by GLM-based and data-driven analysis approaches.

Within participants, we standardized punishment selections across all trials. For each subject, we then calculated correlation coefficients between trial-wise punishment selections and component time course values for time points 3–6 (corresponding to the early task phase in which the punishment selection was made). Thus, for each subject we obtained 72 correlation coefficients (18 components with 4 time points each). We conducted one sample *t*-tests to compare these coefficients across subjects, using a *P*-value threshold of 0.05 Bonferroni corrected for 72 comparisons ($P < 0.0007$). As this analysis was conducted complementary to the main analysis of comparing time courses between conditions, we did not additionally correct for the comparisons between conditions. As noted above, any positive findings from this analysis need to be interpreted with caution.

To test the reproducibility of the results and their dependence on the AROMA motion correction, we repeated the above analyses without running AROMA.

Results

Task-related, condition-unspecific components

Task-related components, which were not differentially modulated during angry and neutral trials, included components spanning mostly sensory and motor areas (Figure 2), thus likely related to task-events as stimulus presentation and button presses. For example, component 2 included mostly auditory cortex and its activity peaked following task offset. This

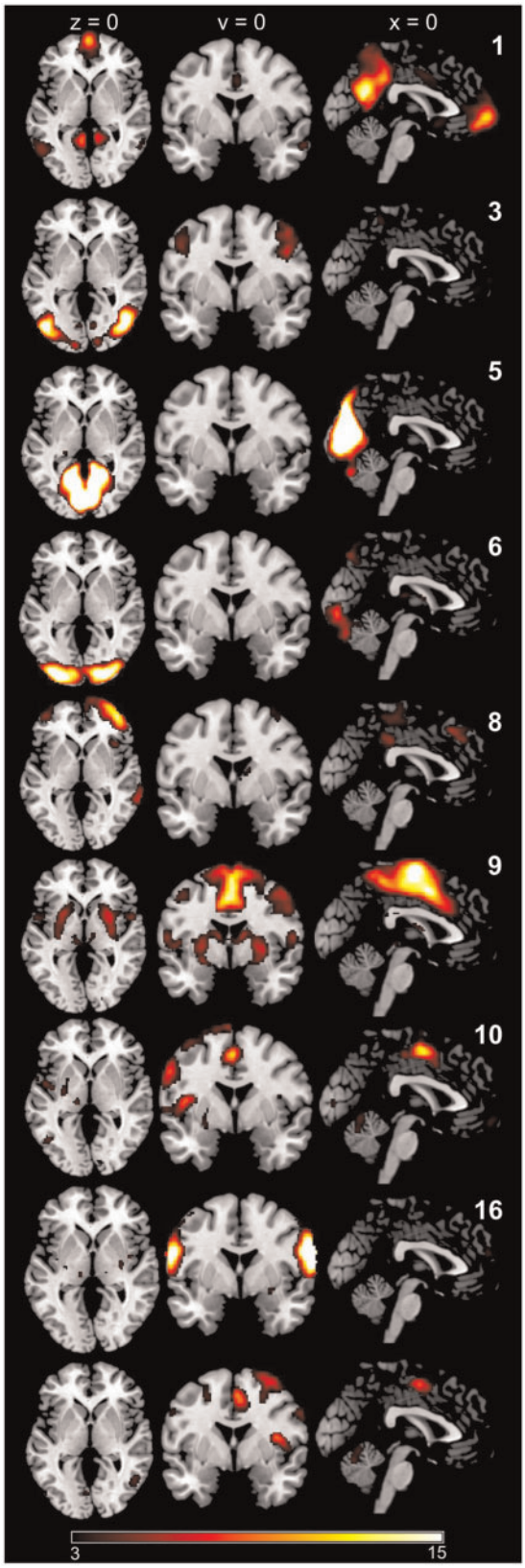

Fig. 2. Component maps. Spatial maps are shown for task-related, condition-unspecific components (excluding those shown in Figure 3). Spatial maps are thresholded at $z > 3.0$.

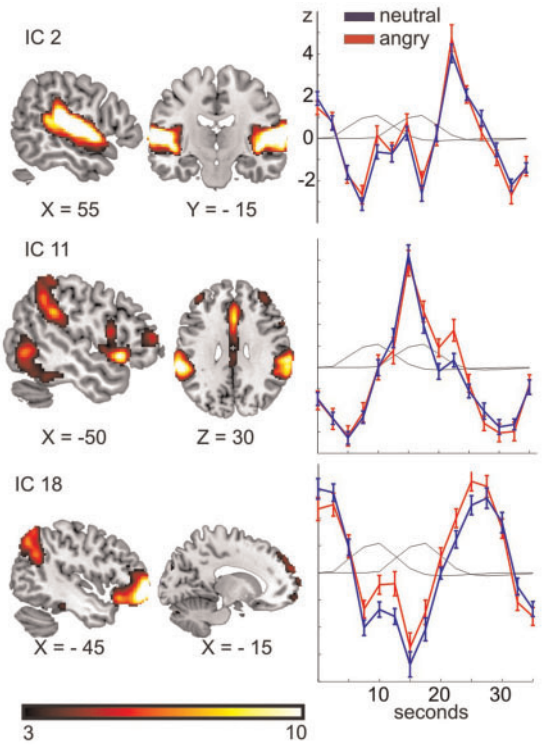

Fig. 3. Condition-unspecific components. Shown are examples of task-related, condition-unspecific components with their spatial maps on the left and the corresponding mean time-courses on the right. Error bars denote standard errors of the mean.

component was therefore likely related to the presentation of the punishment sound (Figure 3). Other components represented activation of the visual system with a time course corresponding to the presentation of visual stimuli in decision and outcome phase (components 5 and 6) and the motor system, corresponding to the button presses during decision phase and reaction time task (component 9).

During the late task phase, increased activity was observed in a network including bilateral inferior and middle frontal gyrus, anterior insula, the cingulate gyrus ranging from anterior to posterior areas, inferior parietal lobule and middle temporal gyrus (component 11; Figure 3). Again, this component did not differentiate between conditions.

Effects of angry *vs* neutral facial expressions on component time courses

In the GLM-analysis, comparing angry > neutral trials showed increased activity during the decision phase in mPFC, inferior frontal gyri and middle/superior temporal gyri. For the outcome phase, this contrast showed increased activity in left temporal pole, middle temporal gyrus and precentral gyrus, the right inferior frontal gyrus, fusiform gyrus, superior parietal lobule, and thalamus (Beyer *et al.*, 2015). Based on the following results, we propose that these activation patterns are related to modulation of several distinct neural systems. We found components that specifically differentiate between task conditions in late task stages, as well as some which show significant differences across the entire task time course.

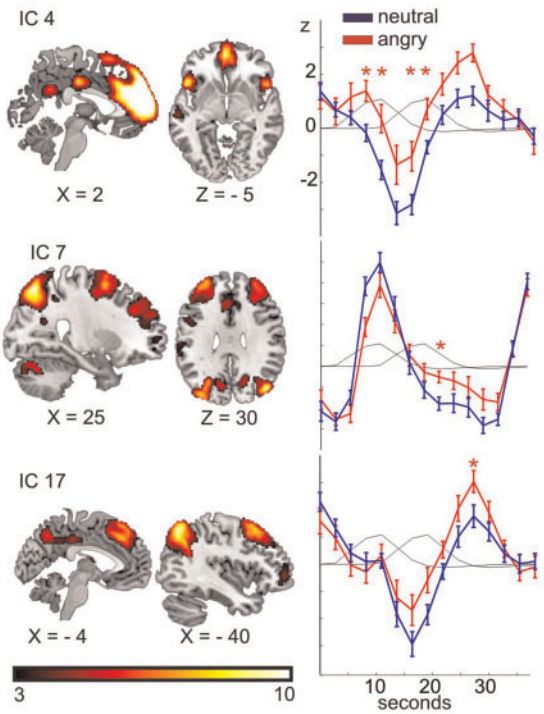

Fig. 4. Task-negative components. Shown are spatial maps of components with decreased activity during the task and significant differences between angry and neutral trials. Mean time-courses for each component are shown on the right, error bars denote standard errors of the mean.

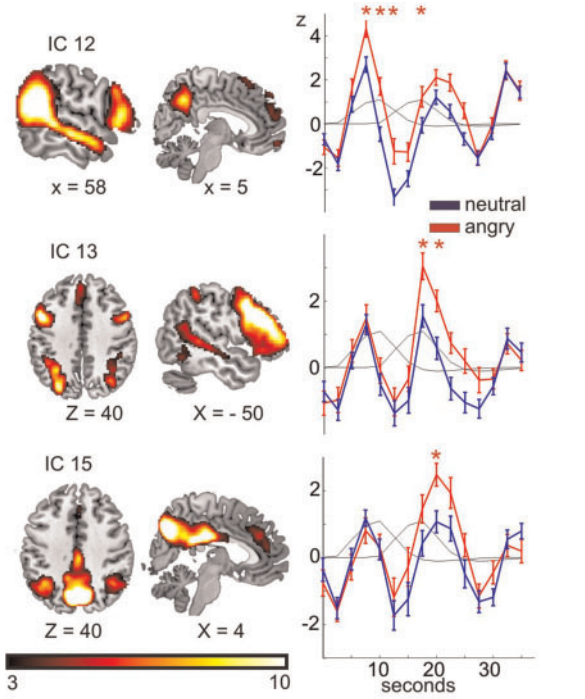

Fig. 5. Task-positive components. Shown are spatial maps of components with increased activity during the task and significant differences between angry and neutral trials. Mean time-courses for each component are shown on the right, error bars denote standard errors of the mean.

Overall, six components showed differential activity during neutral and angry trials, reflected in significant differences between component time courses for angry and neutral trials.

Component 4 showed strong overlap with effects observed using the GLM model, both for the decision and outcome phases. As can be seen in Figure 4, it showed a decrease in activity during neutral trials, but was less modulated by angry trials. The difference between angry and neutral regressors was significant for both early and late task stages (time points 4, 5, 7 and 8). This component included a large area in medial prefrontal cortex ranging from ventral to dorsal areas, as well as bilateral inferior frontal cortex, bilateral middle temporal gyrus, left angular gyrus, the right caudate and two clusters in the cingulate gyrus (Figure 4).

Also for component 12, the difference between angry and neutral trials was significant for both early and late task stages (time points 4, 5, 6 and 8). This component included areas in medial and superior frontal gyrus, bilateral medial and superior temporal gyrus, right middle to inferior frontal gyrus and the precuneus (Figure 5). Thus, these areas partially correspond to those observed for the angry > neutral contrast for the decision phase in the GLM-analysis while for example the precuneus was not observed in the GLM analysis. The time-course for this component however, suggests that this neural system is not exclusively involved in the decision phase, but during later task stages as well (Figure 5).

A network including bilateral middle and superior frontal gyrus, anterior cingulate gyrus, inferior parietal lobule, precuneus, middle temporal gyrus and left inferior frontal gyrus showed increased activity during the early task phase (component 7; Figure 3). This component showed a decrease in activity during later task stages, which was reduced for angry trials.

For component 13, the difference between angry and neutral trials was significant during later task stages (time points 8 and 9). This component included a medial frontal area including part of the supplementary motor area, bilateral areas of middle to inferior frontal gyrus, inferior parietal lobule bilaterally, left superior and middle temporal gyrus, left precuneus and thalamic areas (Figure 5). This component showed increased activity during the late stages of angry trials.

The time-course of component 15 showed a significant difference between angry and neutral trials during late task stages (time point 9). This component spanned posterior cingulate gyrus bilaterally, including most of the precuneus, as well as inferior parietal lobule, superior temporal gyrus and a cluster in the anterior cingulate cortex (Figure 5). Activity in this component showed an increase mainly following the task offset in angry trials. Thus, while being clearly task-related and condition specific, this component's time course did not fit the modelled task time course and consequently includes mainly areas not observed in the GLM analysis.

Similarly, component 17 showed a difference between angry and neutral regressors in the late phase of the task (time point 11). As can be seen in Figure 4, this component showed a task-related decrease in activity, which was stronger in neutral trials, followed by an increase in activity after task offset. The component included areas in bilateral angular gyrus, superior frontal gyrus, medial frontal gyrus and precuneus.

These results suggest that the areas observed in the GLM analysis belong to several distinct and distributed neural systems that are recruited throughout the task. Importantly, the ICA results do not suggest specific modulation of neural systems by task condition during the decision phase. Rather, effects found early on during the task persist during later task stages as well.

Correlations between aggressive behaviour and component time courses did not reach significance for any of the components. Thus, trial-to-trial variability of punishment selections was not consistently related to any of the components' BOLD magnitude time-courses across participants.

To allow for comparison with time courses obtained with the standard GLM analysis, we used the same GLM design as described in Beyer *et al.* (2015). For three key regions of interest from the original angry > neutral contrast reported in Beyer *et al.* (2015) we created 5 mm ROIs around the peak coordinates of the original analysis (mPFC [−2, 44, 40], left superior temporal gyrus [−44, −56, 22], left inferior frontal gyrus [−56, 24, 10]) using the Matlab toolbox marsbar (Brett *et al.*, 2002). We then extracted FIR time courses time locked to trial onset, separately for angry and neutral trials. These are shown in Supplementary material S3.

Re-analysis without AROMA

Without AROMA motion correction on the subject level, the number of non-artefact components was reduced to 11. Consequently, task-related components contained more widespread networks reflecting combinations of the originally found networks. Supplementary material S1 shows task-related networks similar to those reported above, with time courses that are significantly different between conditions. The overall amount of signal variance explained by non-artefact components was lower for this analysis than the ICA run after denoizing with AROMA (see Supplementary material S2). Overall, this re-analysis confirms the reproducibility and robustness of the reported ICA results, as well as the advantage of applying AROMA motion correction prior to the group-level ICA.

Discussion

In this study, we used ICA to explore which distributed neural systems are recruited during different stages of an aggressive interaction, how these networks are affected by angry compared to neutral facial expressions of the task opponent, and whether they are related to aggressive behaviour.

Extending our previous results from standard GLM analyses (Beyer *et al.*, 2015), this approach revealed several task-related networks, some of which were differentially affected by angry and neutral trials.

Modulation of task-positive effects by experimental condition

In angry trials, a fronto-temporal network was recruited during early stages of the task, when participants saw their opponent make his punishment selection and chose a punishment for their opponent (component 12). This network showed a second peak of activity following the late task stages, possibly reflecting processing of the outcome phase information or of the entire trial. The areas constituting this component are mostly part of the mentalizing network (Lieberman, 2007), such as mPFC, middle and superior temporal gyrus, inferior frontal gyrus and precuneus. Thus, it appears that the same cognitive and affective processes evoked by videos of an angry-looking opponent at the beginning of a trial were re-activated following the outcome phase, once participants could integrate the opponent's facial expression with his punishment selection. A similar network was recruited during the outcome phase of angry trials, when participants were informed about their opponent's selection and whether they won or lost (component 13). Thus, these networks reflect the degree to which different task stages were relevant for the social interaction: they were sensitive to the facial expression of the opponent, and they showed two peaks, one during the decision phase and one during the outcome phase.

At the end of angry trials, increased activity was found in a network of anterior cingulate cortex, precuneus, superior temporal gyrus and inferior parietal lobule (component 15). These regions are associated with integration of perception and action (Rizzolatti *et al.*, 2006), visual perspective taking (Schurz *et al.*, 2013) and cognitive control (Shenhav *et al.*, 2013). Possibly, this component reflects the planning of future actions based on the preceding interaction with an angry opponent.

Modulation of task-negative effects by experimental condition

In component 4 (spanning mPFC, IFG, posterior cingulate and STS), we found a negative component time-course, which was more strongly modulated in neutral than angry trials. Thus, while in neutral trials this network showed a decrease in activity, it remained closer to the baseline level in angry trials. It showed an additional increase in activity following angry trials. Thus, baseline activity (or what would be used as baseline in a GLM, i.e. the beginning of the inter-trial interval) in this component was modulated by the preceding trial. The dorsal mPFC and posterior cingulate have been shown to be highly similar in connectivity profiles and functional reactivity, and are related to social cognition and memory processes (Amft *et al.*, 2015). IFG and STS similarly are associated with social cognitive processes as understanding the actions and intentions of someone else (Lieberman, 2007).

Similarly, a network of medial and superior frontal gyrus, angular gyrus and precuneus (component 17) showed a condition-specific peak following the offset of angry trials. This finding indicates the occurrence of mentalizing processes following interactions with a threatening opponent.

Besides being part of the mentalizing network, mPFC, precuneus and superior temporal areas are typically associated with the so-called default mode network, a group of brain areas showing decreased activity in many cognitive tasks, but increased activity in tasks requiring social information processing (Schilbach *et al.*, 2008). Activity in the default mode network during rest has been related to the ease with which participants take the mental perspective of another person and self-report social skills (Spunt *et al.*, 2015). Thus it appears that the decision phase and reaction time task of the TAP principally constitute a (non-social) cognitive task leading to the typical observation of decreased activity in the default mode network. However, in angry trials, this effect is weakened, possibly reflecting a shift from cognitive processing of the immediate task demands towards reflection on the social aspects of the interaction with the opponent.

The default mode network is typically treated as a fixed set of brain regions that show similar patterns of task-related deactivation (Esposito *et al.*, 2006). The exact regions included may differ between studies though and differentiations between core networks and extended networks based on activity and connectivity profiles have been proposed (Amft *et al.*, 2015; M. Göttlich *et al.*, submitted for publication). However, our results suggest that in the context of a social interaction, the DMN splits up into different sub-networks, with partially overlapping spatial maps, but different temporal dynamics. Component 15 spatially corresponds to typically described DMN-regions

observed in resting state (Leech *et al.*, 2011) and shows an increase in activity at the end of angry trials, in correspondence with the idea that this component reflects internally directed thought (Leech *et al.*, 2011). Component 4, on the other hand shows the task-related decrease in activity typical for the DMN, but involves a larger medial prefrontal cluster, whereas the broad PCC cluster typically associated with the DMN is split into two distinct, smaller clusters in the PCC. Component 17 similarly shows the DMN-typical task-related decrease in activity and includes DMN-related areas as mPFC and angular gyrus, but also superior frontal gyrus.

Thus, while with the current study it is not possible to explicitly differentiate between mentalizing and default mode networks, functions associated with both networks, such as contemplation of internal states of the self or other people (Uddin *et al.*, 2007), pose the most plausible explanation for their increased activity during anger trials.

Comparison of ICA and GLM approaches

In a GLM-analysis, each voxel is treated as an independent entity, with separate regression analyses conducted for all voxels. As such, this approach is blind to a priori information about how cognitive processes should be represented in the brain, i.e. the activation of neural systems. Furthermore, each voxel is assigned one coefficient based on the fit of the model to this voxel's time course. This may be problematic in some cases, as different cognitive processes may affect overlapping neural systems, i.e., the activity level of a given voxel may be driven by distinct, potentially counteracting neural processes. ICA, on the other hand, enters all voxels simultaneously into the model estimation and uses spatial information to define neural systems involved in a given cognitive process. Practically, one clear advantage of the ICA approach over GLM analysis is the differentiation between networks that show similar difference values in the same region. For example, both components 4 and 12 showed higher activity in angry than neutral trials in mPFC during the decision phase. However, in component 4 this was due to stronger reduction of activity in neutral trials, in component 12 angry trials produced a stronger task-positive effect. In a GLM analysis, comparing angry *vs* neutral contrast values on the voxel level would produce one positive cluster, combining the results of these two different neural networks, as we found in the GLM analysis conducted previously (Beyer *et al.*, 2015). Similarly, comparing angry against neutral outcome phases in the GLM analysis revealed large, unspecific effects across wide areas of the brain (Beyer *et al.*, 2015). Our current results suggest that this effect may indeed reflect a combination of the effects represented in components 4, 12 and 13, all of which showed increased activity towards the end of angry trials.

Effects involving the regressor in the GLM contrasting angry against neutral trials for the early task stage (the decision phase) were not reflected in our ICA results. We found no condition-specific component time courses that were specifically modulated by early task stages. Rather, the neural systems affected by angry videos in the beginning of a trial showed similar effects during later task stages as well. Furthermore, the time-courses of components 15 and 17, peaking after task offset, were similar to that of the auditory component reflecting the punishment sound. In a GLM analysis, these effects would be difficult to observe as typically, the offset of a task is not explicitly modelled and signal variability related to this would therefore be best explained by the regressor modelling the onset of the punishment sound. The sensitivity of these effects to the experimental condition suggests that cognitive processes related to the social interaction occur during the resting period. Thus, in cases where the time-courses of different neural processes overlap, or are not known a priori, the data-driven analysis approach grants a better understanding of which networks are involved in a given task, and which neural processes contribute to an observed difference value between experimental conditions.

In none of the components did the correlation between trial-wise aggressive behaviour and single-trial time-course values reach significance. Thus, activity in these task-related networks was not consistently related to punishment selection. In our previous analysis we found a trial-wise modulation of dACC activity by punishment selection in angry trials (Beyer *et al.*, 2015). Thus, for such specific investigations of within-subject variability, the GLM approach may be more sensitive. As the concatenation approach for ICA group analysis finds networks that are similar across subjects and trials, neural networks with a high spatial variability across the task may not be detected. Other approaches as investigating subject-specific time courses or using Tensor ICA (Beckmann and Smith, 2005) instead of the concatenation analysis may be more sensitive to variability on the subject level. Tensor ICA however, requires task time courses to be comparable between subjects, which was not the case with our randomized experimental design.

Regardless of the method used, however, the ICA approach by nature is best suited to detect large-scale neural networks, and may be less sensitive for the detection of individual, smaller regions. This limitation can have particular relevance for task-based MRI data with a uniform time course dominating the analysis which, as in the present example, can lead to a relatively low number of identified components.

Another limitation of the ICA approach which has to be considered is the post-hoc nature of interpreting the results. While for resting-state data predictions can be made as to which networks should typically be found, for task-related data this analysis is, at this point, more exploratory. Any interpretation of the functional meaning of observed effects is based on our knowledge stemming from traditional, GLM-based studies showing increased (or decreased) activity in a given brain region during tasks aimed at eliciting certain cognitive or affective processes. Nevertheless, ICA allows for the identification of the neural processes which are recruited during a given task stage. The functional meaning of these neural processes will have to be established by future research implementing ICA on tasks assumed to evoke similar or differential social and cognitive processes.

Conclusions

Using ICA to re-analyse data from a complex social fMRI-paradigm, we compared this approach to classical GLM analysis. We were able to differentiate between several neural systems involved during different stages of the task, which in the GLM analyses resulted in wide-spread, unspecific activation patterns. Furthermore, we found condition-specific component time courses that did not fit the regression model. Thus, in such complex social tasks, this approach can help to understand which neural systems are involved in the task and how their temporal dynamics map onto the task time course. Possibly, findings from an ICA might be used to form optimized GLM regressors in future studies using similar tasks. The ICA was in our case less sensitive to within-subject variability in behaviour than the GLM approach. Thus, to gain an optimal understanding

of task effects on neural processes, combining both analysis strategies would be the recommended approach.

Supplementary data

Supplementary data are available at *SCAN* online.

Funding

This study was supported by the German section of the International Federation of Clinical Neurophysiology, the German Research Foundation (KR3691/5-1) and through intramural funding by the University of Lübeck. FB is supported by the German Research Foundation DFG (BE 5779/1-1). C.F.B. gratefully acknowledges funding from the Wellcome Trust UK Strategic Award [098369/Z/12/Z]. C.F.B. is supported by the Netherlands Organisation for Scientific Research (NWO-Vidi 864-12-003).

Conflict of interest. None declared.

References

Amft, M., Bzdok, D., Laird, A.R., Fox, P.T., Schilbach, L., Eickhoff, S.B. (2015). Definition and characterization of an extended social-affective default network. *Brain Struct Funct* **220**(2), 1031–49.

Beckmann, C.F. (2012). Modelling with independent components. *Neuroimage* **62**(2), 891–901.

Beckmann, C.F., Smith, S.M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans Med Imaging* **23**(2), 137–52.

Beckmann, C.F., Smith, S.M. (2005). Tensorial extensions of independent component analysis for multisubject FMRI analysis. *Neuroimage* **25**(1), 294–311.

Beyer, F., Münte, T.F., Erdmann, C., Krämer, U.M. (2014). Emotional reactivity to threat modulates activity in mentalizing network during aggression. *Soc Cogn Affect Neurosci* **9**(10), 1552–60.

Beyer, F., Münte, T.F., Göttlich, M., Krämer, U.M. (2015). Orbitofrontal cortex reactivity to angry facial expression in a social interaction correlates with aggressive behavior. *Cereb Cortex* **25**(9), 3057–63.

Brett, M., Anton, J.L., Valabregue, R., Poline, J.B. (2002). Region of interest analysis using the MarsBar toolbox for SPM 99. *Neuroimage* **16**(2), S497.

Calhoun, V.D., Adali, T., Pearlson, G.D., Pekar, J.J. (2001). A method for making group inferences from functional MRI data using independent component analysis. *Hum Brain Mapp* **14**(3), 140–51.

Esposito, F., Bertolino, A., Scarabino, T., Latorre, V., Blasi, G., Popolizio, T., Tedeschi, G. (2006). Independent component model of the default-mode brain function: assessing the impact of active thinking. *Brain Res Bull* **70**(4–6), 263–9.

Göttlich, M., Ye, Z., Rodriguez-Fornells, A., Münte, T.F., Krämer, U.M. (2017). Viewing socio-affective stimuli increases connectivity within an extended default mode network. *NeuroImage*, **148** (Supplement C), 8–19.

Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M. (2012). Fsl. *Neuroimage* **62**(2), 782–90.

Kincses, Z.T., Johansen-Berg, H., Tomassini, V., Bosnell, R., Matthews, P.M., Beckmann, C.F. (2008). Model-free characterization of brain functional networks for motor sequence learning using fMRI. *Neuroimage* **39**(4), 1950–8.

Krämer, U.M., Buttner, S., Roth, G., Münte, T.F. (2008). Trait aggressiveness modulates neurophysiological correlates of laboratory-induced reactive aggression in humans. *J Cogn Neurosci* **20**(8), 1464–77.

Krämer, U.M., Jansma, H., Tempelmann, C., Münte, T.F. (2007). Tit-for-tat: the neural basis of reactive aggression. *Neuroimage* **38**(1), 203–11.

Krämer, U.M., Kopyciok, R.P., Richter, S., Münte, T.F. (2009). Oscillatory brain activity related to control mechanisms during laboratory-induced reactive aggression. *Front Behav Neurosci* **3**, 46.

Krämer, U.M., Riba, J., Richter, S., Münte, T.F. (2011). An fMRI study on the role of serotonin in reactive aggression. *PLoS One* **6**(11), e27668.

Leech, R., Kamourieh, S., Beckmann, C.F., Sharp, D.J. (2011). Fractionating the default mode network: distinct contributions of the ventral and dorsal posterior cingulate cortex to cognitive control. *J Neurosci* **31**(9), 3217–24.

Lieberman, M.D. (2007). Social cognitive neuroscience: a review of core processes. *Annu Rev Psychol* **58**, 259–89.

McKeown, M.J., Makeig, S., Brown, G.G., *et al.* (1998). Analysis of fMRI data by blind separation into independent spatial components. *Hum Brain Mapp* **6**(3), 160–88.

Poline, J.B., Brett, M. (2012). The general linear model and fMRI: does love last forever? *Neuroimage* **62**(2), 871–80.

Pruim, R.H., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J.K., Beckmann, C.F. (2015). ICA-AROMA: a robust ICA-based strategy for removing motion artifacts from fMRI data. *Neuroimage* **112**, 267–77.

Rizzolatti, G., Ferrari, P.F., Rozzi, S., Fogassi, L. (2006). The inferior parietal lobule: where action becomes perception. *Novartis Found Symp* **270**, 129–40; discussion 140–125, 164–129.

Schilbach, L., Eickhoff, S.B., Rotarska-Jagiela, A., Fink, G.R., Vogeley, K. (2008). Minds at rest? Social cognition as the default mode of cognizing and its putative relationship to the "default system" of the brain. *Conscious Cogn* **17**(2), 457–67.

Schilbach, L., Timmermans, B., Reddy, V., *et al.* (2013). Toward a second-person neuroscience. *Behav Brain Sci* **36**(4), 393–414.

Schurz, M., Aichhorn, M., Martin, A., Perner, J. (2013). Common brain areas engaged in false belief reasoning and visual perspective taking: a meta-analysis of functional brain imaging studies. *Front Hum Neurosci* **7**, 712.

Scott, G., Hellyer, P.J., Hampshire, A., Leech, R. (2015). Exploring spatiotemporal network transitions in task functional MRI. *Hum Brain Mapp* **36**(4), 1348–64.

Shenhav, A., Botvinick, M.M., Cohen, J.D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* **79**(2), 217–40.

Spunt, R.P., Meyer, M.L., Lieberman, M.D. (2015). The default mode of human brain function primes the intentional stance. *J Cogn Neurosci* **27**(6), 1116–24.

Taylor, S.P. (1967). Aggressive behavior and physiological arousal as a function of provocation and the tendency to inhibit aggression. *J Pers* **35**(2), 297–310.

Uddin, L.Q., Iacoboni, M., Lange, C., Keenan, J.P. (2007). The self and social cognition: the role of cortical midline structures and mirror neurons. *Trends Cogn Sci* **11**(4), 153–7.