Re-submitted to *NeuroImage* on 8th of June 2010

**Dissociating mental states related to doing nothing
by means of fMRI pattern classification**

Simone Kühn^{1,2}, Nils Christian Bodammer³, & Marcel Brass¹

¹Ghent University

Department of Experimental Psychology and Ghent Institute for Metabolic Imaging

Henri Dunantlaan 2, 9000 Gent

Belgium

² Max Planck Institute for Human Cognitive and Brain Sciences,

Department of Psychology,

Stephanstr. 1A, 04103 Leipzig

Germany

³Center of Lifespan Psychology Max Planck Institute for Human Development

Lentzeallee 94, 14195 Berlin

Germany

Corresponding author: Simone Kühn

e-mail: simone.kuhn@ugent.be

Keywords: fMRI, pattern classification, resting- state, voluntary action, voluntary non-action

Abbreviated Title: Dissociating mental states related to doing nothing

Abstract

Most juridical systems recognize intentional non-actions - the failure to render assistance - as intentional acts by regarding them as in principle culpable. This raises the fundamental question whether intentional non-actions can be distinguished from simply not doing anything. Classical GLM analysis on functional magnetic resonance imaging (fMRI) data reveals that not doing anything is associated with resting state brain areas whereas intentionally non-acting is associated with brain activity in left inferior parietal lobe and left dorsal premotor cortex. By means of pattern classification we quantify the accuracy with which we can distinguish these two mental states on the basis of brain activity. In order to identify brain regions that harbour a distributed, overlapping representation of voluntary non-actions and the decision not to act we performed pattern classification on brain areas that did not appear in the GLM contrasts. The prediction rate is not reduced and we show that the prediction relies mostly on brain areas that have been associated with action production and motor imagery as supplementary motor area, right inferior frontal gyrus and right middle temporal area (V5/MT). Hence our data support the implicit assumption of legal practice that voluntary non-action shares important features with overt voluntary action.

Introduction

In daily life we sometimes fail to act in situations where we ought to act. Imagine a situation where someone witnesses a robbery but does not intervene or call the police. The question arises, whether the person explicitly decided not to render assistance or whether he simply did not notice the robbery. From an outward perspective it is difficult - sometimes even impossible - to decide between these two alternatives, because no behaviour is indicative for the difference. However, most legal systems draw a sharp distinction between intentionally not doing something (e.g. failure to render assistance) and unintentionally failing to act. This implicates the assumption that intentionally not doing something is functionally more equivalent to intentionally doing something. But how valid is this assumption from a scientific point of view? Can we distinguish specific brain states that correspond to these two instances of not doing something?

Recent ‘brain-reading’ (Cox & Savoy, 2003) approaches in neuroscience offer a method to quantify the discrimination accuracy of hidden intentional states by means of pattern classification analysis on brain activity. Such approaches have reliably predicted whether a person will choose one task over another (Haynes et al., 2007), which task set a person has currently in mind (Bode & Haynes, 2009) or whether someone is going to press a left or right key (Hampton & O’Doherty, 2007; Soon et al., 2008). However, in these situations the mental state usually results in – specific and measurable - behaviour i.e. pressing a key. Up to now it is not known whether these techniques can decode from brain activity alone, whether a person is simply doing nothing, or intentionally omitting an action.

Therefore the aim of the present study was to explore possible dissociations between voluntarily choosing not to act and not choosing at all. Recently, we demonstrated that voluntary choosing not to act shares functional properties with voluntary actions (Kühn & Brass, 2009; Kühn, Elsner, Prinz, & Brass, 2009; Kühn, Gevers, & Brass, 2009). These findings suggest that brain areas associated with action preparation and motor imagery such as premotor cortex, inferior frontal gyrus and supplementary motor area (SMA) should be involved in choosing not to perform a certain action. In contrast, the mental state of ‘not having decided yet’ is presumably related to the so-called ‘resting-state’ or ‘default mode’ (Raichle et al., 2001) where participants simply do not engage in the task. Mainly four regions are involved in the resting-state: lower precuneus, superior and inferior anterior medial frontal regions, and posterior lateral parietal cortices (Gusnard & Raichle, 2001).

In order to experimentally compare and contrast between a voluntary decision not to act and a situation where participants simply do nothing, we asked our participants to lie in an MR scanner and observe a rotating clock hand (Libet, Gleason, Wright, & Pearl, 1983) which they could set in motion by pressing a button. They were instructed to decide - at a point in time of their own choice - between voluntarily acting and voluntarily not acting while the clock hand rotated for a random interval. Participants were informed that the clock interval would from time to time terminate before they actually could come up with a decision. After each clock interval participants had to indicate whether they decided to act (Go), decided not to act (NoGo) or did not decide at all (No-Decide) (*Figure 1*).

From a third-person perspective, it is difficult to tell whether a participant is intentionally ‘deciding not to act’ or merely idling, because both states look highly similar: the participant does not respond. However, we assume that the brain activity associated with both mental states

should differ and should thus allow us to predict whether participants voluntarily refrained from acting or whether they simply did not make a decision.

Methods

Participants

Eighteen healthy volunteers participated on the basis of informed consent and with ethical committee approval and according to the Declaration of Helsinki. All participants had normal or corrected-to-normal vision and no history of neurological, major medical, or psychiatric disorder. The data of three subjects were excluded from the analysis due to strong movement artifacts. The remaining fifteen subjects (9 women and 6 men; age: mean = 21.6, ranging from 18 to 26) were all right-handed as assessed by a handedness questionnaire (van Strien, 1992; mean score = 9.7).

Behavioural Task

Participants were instructed to freely decide when and whether to execute a simple key press while observing a rotating clock hand (Libet et al., 1983). Participants were told to take their time and to decide between responding with their right index finger (and then executing the response) or not responding during the rotation of the clock hand. They were encouraged to decide equally often to execute or not to execute the response. Each trial started with a clock face on the screen. Participants initiated the trial by pressing the key with their right index finger. After a delay of 500 ms the clock hand appeared at a random position and began to rotate (period 4000 ms). The clock hand stopped rotating after a random interval of 3000 to 6000 ms after the button press. After a blank screen of 2000 - 5000 ms (varied in steps of 500 ms) a request screen was

presented. Participants had to select whether they decided to respond (depicted as a hand pressing a button), decided not to respond (depicted as a crossed out hand pressing a button) or did not decide yet (depicted as a cloud) during the clock presentation. Additionally there was a response option to indicate that participants were not sure (depicted as a question mark). Participants selected their response to the request screen using either index or middle finger of both hands. The key mapping varied randomly throughout the experiment in order to prevent advance response preparation. After a blank screen of 2 s the next trial started. After each run response feedback was presented to the participants. The experiment consisted of nine blocks of 27 trials and lasted approximately 60 min. Before scanning the subjects received one practice block outside the scanner.

Scanning Procedure

Images were collected with a 3T MagnetomTrio MRI scanner system (Siemens, Erlangen, Germany) using an 8-channel radiofrequency head coil. High-resolution anatomical images were acquired using a T1-weighted 3D MPRAGE sequence (TR = 2530ms, TE = 2.58ms, TI = 1100ms, acquisition matrix = 176×256×256, sagittal FOV = 220 mm, flip angle = 7°, voxel size = 0.9×0.86×0.86 mm³). Whole brain functional images were collected using a T2*-weighted EPI sequence sensitive to BOLD contrast (TR = 2000ms, TE = 35ms, image matrix = 64×64, FOV = 224 mm, flip angle = 80°, slice thickness = 3.0 mm, distance factor = 17%, voxel size 3.5×3.5×3 mm³, 30 axial slices).

fMRI Data Pre-processing

The fMRI data were preprocessed using statistical parametric mapping using the SPM5 software (Wellcome Department of Cognitive Neurology, London, UK). The first 4 volumes of all EPI

series were excluded from the analysis to allow the magnetisation to approach dynamic equilibrium. Data processing started with slice time correction and realignment of the EPI datasets. A mean image for all EPI volumes was created, to which individual volumes were spatially realigned by rigid body transformations. The high resolution structural image was co-registered with the mean image of the EPI series. The structural image was then normalised to the Montreal Neurological Institute (MNI) template, and the normalisation parameters were applied to the EPI images. A commonly applied filter of 8 mm FWHM (full-width at half maximum) was used. Low-frequency drifts in the time domain were removed by modelling them by a set of discrete cosine functions to which a cut-off of 128 s was applied.

GLM analysis

The main event of interest was set to the time point of the average response time in the Go condition calculated from the start of the clock. Vectors containing the event onsets (Go, NoGo, No-Decision, request screen) were convolved with the canonical haemodynamic response function (HRF) to form the main regressors in the design matrix (the regression model). The vectors were also convolved with the temporal and dispersion derivatives and the resulting vectors were entered into the model. Additionally we entered the movement parameters of the realignment procedure as a regressor into the model. The statistical parameter estimates were computed separately for each voxel for all columns in the design matrix. Contrast images were constructed from each individual subject to compare the relevant parameter estimates for the regressors containing the canonical HRF. On the group-level a one-sample t-test was performed for each voxel of the contrast images. The resulting statistical values were thresholded with a lenient threshold at $p < 0.001$ ($z > 3.09$, uncorrected) with a volume greater than 350 mm^3 (10 adjacent voxels).

Pattern classification

For pattern classification linear support vector machines (SVMs) were trained on subsets of the normalised functional data time series (Pereira, Mitchell, & Botvinick, 2009). To reduce the dimension of feature space, only those voxels which contain grey-matter with high probability were selected. To further reduce the size of data principal component analysis (PCA) was used to express the time series of functional image data by vectors containing the eigenvalues assigned to the principal component images. Whenever the data dimensionality (voxels) exceeds the number of data points (scans) one can represent the data in a space of reduced dimensionality – spanned by the principle components – without loss of information (Ringnér, 2008). The coordinates of the training and the test data were linearly transformed into the vector space spanned by the full set of principal components which now form the new vector basis of feature space (procedure described in Mourao-Miranda et al., 2005). Then linear SVMs (Spider Toolbox, <http://www.kyb.tuebingen.mpg.de/bs/people/spider/main.html> by Weston, Elisseeff, Bakir, & Sinz) were trained and tested with the projected data. We used the SVM classifier to assess how much information is encoded in the brain pattern that discriminates between time intervals in which participants voluntarily decided not to act (NoGo) and intervals in which they did not decide not to act (No-Decision). The onset times according to which we assigned the subsequent two image volumes to the respective condition was the clock-start time for each trial. Additionally we introduced a lag of 4s to account for the delay of the BOLD response. We then assigned the pattern vectors of all image volumes acquired during $k-1$ imaging runs ($k = 9$ runs) to a training data set that was used to train a linear SVM with fixed regularisation parameter $C = 1$ to correctly identify response patterns related to NoGo and No-Decision. The amount of information present could then be assessed by examining how well the presence or absence of intentions during the

remaining independent test data were classified. In total, the training and test procedure was repeated 9 times, each time with a different run assigned as test data set (k-fold cross-validation). The class labels assigned to the test trials by the classifier were compared to the experimentally obtained class labels to calculate the accuracy of the classification, i.e. the correct prediction rate. Additionally, we computed the ability to accurately recall information during the test procedure according to information processing theory. This recall measure is defined as the proportion of the number of trials that are correctly assigned to a certain class by the classifier and the number of trials that actually belong to that respective class according to the participants' response to the request screen. Finally the 'average recall' – defined as the mean of the recall rates for both classes – was used because it is insensitive to size differences between the classes.

Permutation tests

In order to assess the reliability of the classification we compared the classification results to guessing level. For this purpose we performed permutation tests in which we randomly permuted the class labels of the training data. Cross-validation was then performed on the permuted training set. The permutation was repeated 1000 times for each participant's dataset. The resulting estimates were used to determine the individual mean guessing level at the 95% confidence interval. We assume that the classifier learned reliably from the data when the prediction rates obtained by the classifier trained on the real class labelling exceeds the 95% confidence interval of the classifier trained on randomly relabelled class labels. We report p -values that represent the probability of making an error when rejecting the null hypothesis that the classifier trained on the real class labels did not learn to classify.

Spatial discriminance maps

With each trained linear SVM an optimal class-separating hyperplane was specified in PCA space by the weight vector w , which is a normal vector to this hyperplane and, thus, represents the direction of the highest separability. When the image volumes of all trials are used for training the absolute value of each entry in w can be interpreted as a measure for the information the corresponding principal component contributes to the classification. After projecting the w -vector given by its PCA eigenvalues back into voxel space the resulting voxel values were computed to visualize the information each voxel contributes to the classification performance. According to the two classes (NoGo: +1; No-Decision: -1) a positive value in these maps – which we term spatial discriminance maps (SDM) (Wang, Childress, Wang & Detre, 2007) – indicates that this voxel contributes to the overall classification performance and its signal intensity on average was higher during class +1 compared to class -1. Since the subjects' images have previously been normalized to the MNI template, we performed a second-level analysis of all subjects' SDMs containing absolute values of classification performance (FWE = 0.05, $k = 10$). The result indicates how much each individual voxel on average contributes to the classification across all subjects. First we performed the classification on the whole brain grey matter and then on the whole brain grey matter excluding voxels that were significant in the GLM analysis ($p < 0.001$, with a volume greater than 10 adjacent voxels). We applied cross-validation and permutation tests for each subject individually and obtained average recall measures for both correct and random labelling. We then calculated the mean values over all subjects for both the average recall and the p-value.

Results

Behavioural data

The mean time until a response occurred in the Go condition was on average 2.92 s after clock rotation onset. The distribution of Go, NoGo and No-Decision trials did not differ significantly (32.7%, 33.3%, 31.5%; paired t-tests: $p > 0.36$). Likewise we obtained on average an equal amount of NoGo and No-Decision trials (75 NoGo, 71 No-Decision; $t(14) = .65$, $p = 0.53$). The average trial duration was not significantly different (5.08s NoGo, 5.00s No-Decision, $t(14) = 1.25$, $p = 0.23$). In order to exclude that participants use simple alternation strategies to achieve an equal distribution we calculated the mean Random Number Generation (RNG) (Evans, 1978). This measure considers the randomness of the sequence, namely the dependency between one choice and the next. Our results yielded an average RNG of 0.753. Comparing this to randomly produced sequences of three equally distributed response options no significant difference can be detected (RNG = 0.752, $t(14) = 0.53$, $p = 0.60$).

fMRI data

GLM analysis

When comparing NoGo trials with No-Decision trials we found significant activation in left inferior parietal lobule (IPL) extending into the left superior temporal gyrus (STG) and the left dorsal premotor cortex (dPMC) (*Table 1A, Figure 2A*). In line with our predictions we found higher activation in the so-called resting state network in the reverse contrast (No-Decision > NoGo): paracingulate gyrus, lower precuneus, right superior frontal gyrus and right IPL (*Table 1B, Figure 2B*).

Whole-brain classification

We applied pattern classification on all grey-matter voxels of the brain in order to quantify the accuracy with which the data allows to distinguish between the intention not to do something (NoGo) and not intending anything (No-Decision). The average prediction accuracy (average recall) over all participants was with 72.5 % significantly above chance ($p < 0.002$) when tested against the distribution of average recall rates obtained through random labelling in permutation tests.

Since it has lately been an issue of debate whether it is beneficial to use unsmoothed imaging data when using pattern classification methods, we repeated the whole brain analysis on unsmoothed data. The results were highly similar although the average recall rate was slightly lower 71.8% ($p < 0.002$). When comparing percent correct prediction instead of average recall rates on smoothed (73.8%) and unsmoothed data (73.2%) the tendency was the same, namely slightly better prediction rates for smoothed compared to unsmoothed data.

Classification excluding brain regions that were significant in the GLM analysis

Since whole brain classification methods do not go far beyond GLM analyses when classification is obtained from areas that likewise appear in the GLM contrasts we used a brain mask excluding all brain areas that did show up in a GLM analysis, those resulting from the contrast No-Decision > NoGo and the reverse contrast NoGo > No-Decision. The average prediction accuracy (average recall) over all participants was still significantly above chance (72.84 %, $p < 0.002$), indicating that classification performance is not hampered by the exclusion of brain areas which seem to display the most important difference between the two conditions in a GLM analysis.

Furthermore focussing on the resulting SDMs allows us to search for brain regions that harbour a distributed and overlapping representation of voluntary non-actions and no-decision not to act. Second-level analysis on SDMs of individual subjects revealed high classification properties in

brain areas that have been shown to be active during motor imagery as right inferior frontal gyrus (IFG, pars opercularis/triangularis, BA 44/45; MNI coordinates: 53, 32, 0), SMA (BA 6; 11, 7, 60); as well as right motion-sensitive middle temporal area (V5/MT, BA 19; 39, -67, -7), right anterior insula/ operculum (BA 13; 35, 18, 0) and bilateral precuneus (BA 7; 21, -70, 39; -14 -81, 39) (*Figure 3*). In order to ensure that those brain areas indeed display no overall signal difference we performed percent signal change analyses (in a window of 6-8s with a sphere of 6 mm radius around the peak voxel) for the brain regions that appeared to be highly predictive in the pattern classification analysis. None of these regions of interest, except bilateral precuneus ($t(14) = -2.21$, $p < 0.05$) showed a significant signal change difference ($p > .33$).

Discussion

Although it is difficult to judge from a third-person perspective whether someone is intentionally doing nothing or merely idling, it is quite obvious from the first person perspective. Our results demonstrate that fMRI data allow us distinguish between those inner states of mind. By means of pattern classification on the whole brain we can quantify the prediction and reach 72% accuracy.

Distinct representation of “doing nothing”

By means of classical GLM analysis we determined brain areas that distinguish between voluntary non-action (NoGo) and not having decided yet (No-Decision) in an on-and-off-fashion. Brain areas that displayed higher activation levels during intentionally not acting compared to No-Decisions are left dPMC cortex and left IPL (extending into superior temporal gyrus). The premotor cortex is usually assumed to be involved in processes related to motor preparation (Freund, 1990; Halsband, Ito, Tanji, & Freund, 1993). Dorsal premotor cortex in particular has

been associated with the selection of action (Cavina-Pratesi et al., 2006; O'Shea, Johansen-Berg, Trief, Göbel, & Rushworth, 2007; Toni et al., 2002) and the generation of motor plans (de Lange, Hagoort, & Toni, 2005). The inferior parietal lobe (IPL) has been associated with action awareness and agency in the context of voluntary action. IPL activation has e.g. been shown to increase when participants attend to their intention to move (Lau et al., 2004) or their own actions (Farrer et al., 2008). Here, interestingly, we find the left IPL associated with voluntary non-action. Moreover activity in left IPL has been associated with internally monitoring of self-generated compared to externally generated movements (Ogawa & Inui, 2007). Furthermore Desmurget and colleagues were able to elicit the desire to move by electrical stimulation of IPL in humans (Desmurget et al., 2009). The authors argue that parietal cortex generates and monitors movement intention and awareness (Desmurget & Sirigu, 2009).

In sum, brain areas that have been associated with motor intentions and monitoring of actions in the sense of action awareness and agency ascription are involved in voluntary non-actions although they do not involve any overt movement. This strongly supports our assumption that voluntary non-actions bear high resemblance to voluntary actions (Kühn & Brass, 2009; Kühn et al., 2009; Kühn, Gevers, & Brass, 2009).

Consistent with our predictions simply idling or 'not deciding' activated brain areas in an-on-and-off-fashion that have been associated with 'resting-state' (Gusnard & Raichle, 2001; Raichle et al., 2001). This so-called 'default network' comprises paracingulate cortex, lower precuneus, right superior frontal gyrus and right IPL. While the specific functional role of this network is not clear, it becomes activated when participants are not involved in an explicit task. Mason and colleagues (Mason et al., 2007) e.g. demonstrated that mind wandering during task performance is associated with activity in this network. Previous studies on resting state mostly relied on data where a fixation baseline condition or an implicit baseline is compared to a task with a high

central executive demand (Gusnard & Raichle, 2001; McKiernan, Kaufmann, Kucera-Thompson, & Binder, 2003). In contrast to this the present study is to our knowledge the first that compares resting state to a perceptually and motorically equivalent condition, namely a condition where participants do not do anything that poses an alternative to exploring resting state by means of functional connectivity (Grecius, Krasnow, Reiss, & Menon, 2003).

Distributed and overlapping representations of “doing nothing”

The classical GLM approach solely reveals voxels that differ in the height of activity between conditions. It has been suggested that this univariate GLM approach may be suboptimal for fMRI data analysis because it treats each voxel's time series as an independent event ignoring the abundant spatial correlations of fMRI data, so that it can not take advantage of the fact that several brain regions may respond simultaneously. Spatial smoothing in GLM approaches can partly take the spatial correlations into account, but it can only consider those existing in a constrained neighbourhood (Mourao-Miranda et al., 2005). Another problem the univariate GLM approach does not account for is that the actual shape of the hemodynamic response function (HRF) may differ from the canonical one in different voxels. This problem is circumvented by pattern classification approaches that involve no assumptions about the shape of the HRF.

Beyond those advantages pattern classification has been said to be useful in order to answer three different types of questions, namely whether there is information in the brain data about the variable of interest (pattern discrimination), where this information can be found in the brain (pattern localization) and how this information is encoded (pattern characterization) (Peirera, Mitchell, & Botvinick, 2009). We made use of the classifiers discrimination and localization properties. Moreover since it has been suggested that pattern classification allows the detection of more subtle representational differences between brain sets we excluded brain areas that showed

a distinct on-and-off activation difference in the GLM (Haxby et al, 2001; Haynes & Rees, 2006) and performed pattern classification analysis on the remaining voxels. The accuracy of prediction on those remaining voxels was similar to the accuracy achieved on the whole brain data. This reveals that motor intentions are more broadly distributed across the brain than a standard GLM analysis would suggest.

The resulting brain areas that showed high classification properties but no overall activity difference between NoGo and No-Decision trials were areas that have been shown to be related to motor imagery as right IFG and SMA (Jeannerod & Decety, 1995; Kasess et al., 2008; Stephan et al., 1995). In a study that decoded task-sets that participants were prepared for left IFG has been implicated (Bode & Haynes, 2009). The authors suggest that it might house the representation of the task-set. Interestingly the IFG – in particular the pars opercularis as part of the so-called ‘mirror neuron system’ – has been suggested to contain a motor representation of hand movements (Iacoboni et al., 1999; Rizzolatti & Craighero, 2004). Since the activity is right not left lateralized we conclude that the activity does not result from verbalization processes. The fact that we find no overall activity difference in these brain areas but can still decode the type of “doing nothing” implies that voluntary non-action as well as idling both evoke a distributed though differential pattern of activity in those areas.

Adjacent to the activation in right IFG we found another area with high across-subject classification properties in right anterior insula/ operculum (AIC). The AIC has been associated with self-recognition (Devue et al., 2007) and agency (although our coordinates are slightly more anterior; Farrer & Frith, 2002). As Craig (2009) suggests in his review the AIC plays a fundamental role in the “awareness of the immediate moment” (p. 63) and can be considered as a potential neural correlate of consciousness. In a study that classified reward-related decision making AIC has also been shown to be an informative region to predict the outcome of choices

from (Hampton & O'Doherty, 2007). These findings imply that the AIC codes the difference in awareness between intentional and passive states of non-action.

Another area that shares this property is the right V5/MT. Although V5/MT is classically observed during visual presentation of motion (Tootell et al., 1995), V5/MT activity has been shown to occur in the absence of moving stimuli as well; for example when participants perceive motion in illusions, when they imagine moving stimuli (Goebel et al., 1998; Shelton & Pippitt, 2006; Slotnick, Thompson, & Kosslyn, 2005) or when they access knowledge of action words (Kable, Lease-Spellmeyer, & Chatterjee, 2002). Therefore the distributed pattern in V5/MT could imply that voluntary non-actions involve a different kind of motion representation compared to simply not doing anything. The precuneus also revealed classification properties, but also a ROI analysis showed a significant difference in activation height between conditions. This region has been demonstrated to contain predictive information about the outcome of a free choice situation after the selection of the action (Soon et al., 2008). Since precuneus has been implicated in episodic memory tasks (Fletcher et al., 1995) it might comprise information about past trials in order to keep the instructed balance between the different action (and non-action) alternatives.

In order to explore the influence of smoothing on pattern classification we repeated the whole brain classification on unsmoothed data and found basically the same results as on smoothed data. This argues against the assumption that pattern classification techniques allow to differentiate patterns of activity at a finer scale than conventional univariate analysis (Haynes & Rees, 2005; Kamitani and Tong, 2005) and is in line with recent reports by Op de Beeck (2010a, 2010b) arguing against this so called “hyperacuity” hypothesis. Nevertheless the effects of smoothing on pattern classification might depend on the type of experiment and the specific brain networks involved.

Further research is needed in order to exploit the properties of classifiers to provide information about how the information of interest is encoded (pattern characterization). One option to explore this would be to disentangle the representation of different kinds of non-actions as e.g. thinking about not pressing the button with the right compared to the left index finger. This could yield potentially interesting insights into the representation of action in the brain.

Practical implications

Lately it has been suggested that combining law and neuroscience might enrich both fields and may lead to a new specialized discipline in the future (Zeki & Goodenough, 2004). As outlined in the introduction, the difference between intentional and non-intentional non-action can have important juridical implications. Intentional non-actions, being omissions of actions for which the agent anticipated their consequences and has engaged in counterfactual thinking while deciding between acting and not acting (Baird & Fugelsang, 2004). Most juridical systems recognize intentional non-action - e.g. the failure to render assistance - as intentional acts by regarding it in principle as culpable. Under the label of ‘crime of omission’ the failure to act can be regarded as an *actus reus* (guilty act). This concept implicitly assumes that human beings can have a representation of an act when they choose not to act and can therefore be accused of *mens rea* (guilty mind) when voluntarily omitting an action. This *mens rea* is often important in legal determinations of guilt and liability. When judges have to determine whether a defendant acted with “intent”, they are often forced to infer the defendant’s state of mind from facts that can only be gathered about actions (O’Hara, 2004). Our approach seems to be a first step into objectifying these inferences. The results presented suggest that the motor simulation and action awareness network is involved in voluntary non-action and the detectable difference between choosing not to act and simply doing nothing, provide first neuroscientific evidence for the appropriateness of

this course of action. Nevertheless it has long been argued that brain imaging evidence is not appropriate as forensic evidence, especially because one has to make an inference based on group studies, which do not necessarily allow conclusions for a single suspect (Mayberg, 1992, 1996). Although this problem is circumvented by means of pattern classification on single subject data, but the ubiquitous problem is, that in order to ascertain mens rea, the defendant's cognition at the time of the crime or action omission itself has to be measured, and cannot easily be inferred after the fact (Martell, 2009).

Conclusion

With the current study we addressed the question whether intentional non-actions can be distinguished from simply not doing anything. Classical GLM analysis data revealed the involvement of different brain areas. Not doing anything was associated with resting state brain areas whereas intentionally non-acting - despite the lack of overt action - was associated with brain activity in left IPL and left dPMC, brain areas that have been associated with motor intentions, action awareness and agency ascription. By means of pattern classification we demonstrated that these two mental states on the basis of brain activity can be reliably distinguished on the basis of brain activity. In order to identify brain regions that house distributed, overlapping representations of voluntary non-actions we performed pattern classification on brain areas that did not appear in the GLM contrasts. The results revealed that the prediction relies mostly on brain areas that have been associated with action production and motor imagery as SMA, right IFG and right V5/MT. Hence our data support the implicit assumption of legal practice that deciding not to act shares important features with overt voluntary action.

Acknowledgement

We would like to thank Patrick Haggard for helpful discussions, Barbara Müller and Heiko Bergmann for their help with handling the participants and Eamonn Walsh for comments on a first draft of the manuscript. This work was supported by a special research fund (BOF, University of Ghent). The first author is a Postdoctoral Fellow of the Research Foundation Flanders (FWO).

References

Baird, AA, & Fugelsang, JA (2004). The emergence of consequential thought: Evidence from neuroscience. *Philosophical Transactions of the Royal Society*, 359, 1797-1804.

Blakemore S-J, Frith CD, Wolpert DM. 2001. The cerebellum is involved in predicting the sensory consequences of action. *NeuroReport*, 12, 1879-1884.

Bode, S., & Haynes, J.D. (2009). Decoding sequential stages of task preparation in the human brain. *NeuroImage*, 45(2), 606-613.

Cavina-Pratesi C, Valyear KF, Culham JC, Köhler S, Obhi SS, Marzi CA, Goodale MA. 2006. Dissociating arbitrary stimulus-response mapping from movement planning during preparatory period: evidence from event-related functional magnetic resonance imaging. *Journal of Neuroscience*, 26, 2704–2713.

Cox DD, Savoy RL. 2003. Functional magnetic resonance imaging (fMRI) “brain reading”: detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage*, 19, 261-270.

Craig AD. 2009. How do you feel – now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, 10, 59-68.

de Lange FP, Hagoort P, Toni I. 2005. Neural topography and content of movement representations. *Journal of Cognitive Neuroscience*, 17(1), 97-112.

Desmurget, M., Reilly, K.T., Richard, N., Szathmari, A., Mottolese, C., & Sirigu, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science*, 324, 811-813

Desmurget, M., & Sirigu, A. (2009). A parietal-premotor network for movement intention and motor awareness. *Trends in Cognitive Sciences*, 13(10), 411-419.

Devue C, Collette F, Balteau E, Degueldre C, Luxen A, Maquet P, Bredart S. 2007. Here I am: The cortical correlates of visual self-recognition. *Brain Research*, 1143, 169-182.

Evans FJ. 1978. Monitoring attention deployment by random number generation: An index to measure subjective randomness. *Bulletin of the Psychonomic Society*, 12, 35-38.

Farrer C, Frey SH, Van Horn JD, Tunik, E, Turk D, Inati S, Grafton ST. 2008. The angular gyrus computes action awareness representations. *Cerebral Cortex*, 18, 254-261.

Farrer C, Frith CD. 2002. Experiencing oneself vs. another person as being the cause of an action: the neural correlates of the experience of agency. *NeuroImage*, 15, 596-603.

Fletcher, PC, Frith, CD, Baker, SC, Shallice, T, Frachowiak, RSJ, & Dolan, RJ (1995). The mind's eye- Precuneus activation in memory-related imagery, *NeuroImage*, 2, 195-200.

Freund HJ. 1990. Premotor area and preparation of movement. *Rev Neurol (Paris)*, 146, 543.

Goebel R, Khorram-Sefat D, Muckli L, Hacker H, Singer W. 1998. The constructive nature of vision: Direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *European Journal of Neuroscience*, 10, 1563-1573.

Greicius MD, Krasnow B, Reiss AL, Menon V. 2003. Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proceedings of the National Academy of Science*, 100, 253.

Gusnard DA, Raichle ME. 2001. Searching for a baseline: functional imaging and the resting human brain. *Nature Neuroscience Reviews*, 2, 685-694.

Halsband U, Ito N, Tanji J, Freund HJ. 1993. The role of premotor cortex and the supplementary motor area in the temporal control of movement in man. *Brain*, 116, 243.

Hampton, A.N., & O'Doherty, J.P. (2007). Decoding the neural substrates of reward-related decision making with functional MRI. *Proceedings of the National Academy of Science*, 104, 1377-1382.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293, 2425-2430.

Haynes, J.D., & Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience*, 8, 686–691.

Haynes JD, & Rees G. 2006. Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7, 523-534.

Haynes J-D, Sakai K, Rees G, Gilbert S, Frith C, Passingham RE. 2007. Reading hidden intentions in the human brain. *Current Biology*, 17, 323-328 (2007).

Iacoboni M, Woods RP, Brass M, Bekkering H, Mazziotta JC, Rizzolatti G. 1999. Cortical mechanisms of human imitation. *Science*, 286, 2526–28.

Jeannerod M, Decety J. 1995. Mental motor imagery: a window into the representational stages of action. *Current Opinion in Neurobiology*, 5, 727–732.

Kable JW, Lease-Spellmeyer J, Chatterjee A. 2002. Neural substrates of action event knowledge. *Journal of Cognitive Neuroscience*, 14(5), 795-805.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, 8, 679–685.

Kasess CH, Windischberger C, Cunnington R, Lanzenberger R, Pezawas L, Moser E. 2008. The suppressive influence of SMA on M1 in motor imagery revealed by fMRI and dynamic causal modeling. *NeuroImage*, 40, 828-837.

Kühn S, Brass M. 2009. When doing nothing is an option: The neural correlates of deciding whether to act or not. *NeuroImage*, 46(4), 1187-1193.

Kühn S, Elsner B, Prinz W, Brass M. 2009. Busy doing nothing: Evidence for non-action-effect binding. *Psychonomic Bulletin Review*, 16(3), 542-549.

Kühn S, Gevers W, Brass M. 2009. The neural correlates of intending not to do something. *Journal of Neurophysiology*, 101(4), 1913-1920.

Lau HC, Rogers RD, Haggard P, Passingham RE. 2004. Attention to intention. *Science*, 303, 1208–1210.

Libet B, Gleason CA, Wright EW, Pearl DK. 1983. Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain*, 106, 623-642.

Martell, D.A. (2009). Neuroscience and the law: Philosophical differences and practical constraints. *Behavioral Sciences and the Law*, 27, 123-136.

Mason MF, Norton MI, Van Horn JD, Wegner DM, Grafton ST, Macrea CN. 2007. Wandering minds: The default network and stimulus-independent thought. *Science*, 315, 393-395.

McKiernan KA, Kaufman JN, Kucera-Thompson J, Binder JR. 2003. A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *Journal of Cognitive Neuroscience*, 15(3), 394-408.

Mayberg, H.S. (1992). Functional brain scans as evidence in criminal court: An argument for caution. *Journal of Nuclear Medicine*, 33(6), 18-19.

Mayberg, H.S. (1996). Medical-legal inferences from functional neuroimaging evidence. *Seminar in Clinical Neuropsychiatry*, 1(3), 195-201.

Mourao-Miranda J, Bokde ALW, Born C, Hampel H, Stetter M. 2005. Classifying brain states and determining the discriminating activation patterns: Support vector machine on functional MRI data. *NeuroImage*, 28, 980-995.

O'Hara, E.A. (2004). How neuroscience might advance the law. *Philosophical Transactions of the Royal Society*, 359, 1677-1684.

Op de Beeck, H.P. (2010a). Against hyperacuity in brain reading: Spatial smoothing does not hurt multivariate fMRI analyses? *NeuroImage*, 49, 1943-1948.

Op de Beeck, H.P. (2010a). Probing the mysterious underpinnings of multi-voxel fMRI analyses. *NeuroImage*, 50, 567-571.

O'Shea J, Johansen-Berg H, Trief D, Göbel S, Rushworth MFS. 2007. Functionally specific reorganization in human premotor cortex. *Neuron*, 54, 479.

Ogawa K, Inui T. 2007. Lateralization of the posterior parietal cortex for internal monitoring of self- versus externally generated movements. *Journal of Cognitive Neuroscience*, 19(11), 1827-1835.

Pereira F, Mitchell T, Botvinick M. 2009. Machine learning classifiers and fMRI: A tutorial overview. *NeuroImage*, 45 (Suppl), S199-209.

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., & Shulman, G.L. (2001). A default mode of brain function. *Proceedings of the National Academy of Science*, 98(2), 676-682.

Ringnér, M. (2008). What is principal component analysis? *Nature Biotechnology*, 26(3), 303-304.

Rizzolatti G, Craighero L. 2004. The mirror-neuron system. *Annual Reviews Neuroscience*, 27, 169-192.

Shelton AL, Pippitt HA. 2006. Motion in the mind's eye: Comparing mental and visual rotation. *Cognitive, Affective & Behavioral Neuroscience*, 6(4), 323-332.

Slotnick SD, Thompson WL, Kosslyn SM. 2005. Visual mental imagery induces retinotopically organized activation of early visual areas. *Cerebral Cortex*, 15, 1570-1583.

Soon CS, Brass M, Heinze H-J, Haynes J-D. 2008. Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, 11(5), 543-545.

Stephan KM, Fink GR, Passingham RE, Silbersweig D, Beballos Baumann AO, Frith CD, Frackowiak RS. 1995. Functional anatomy of the mental representation of upper extremity movements in healthy subjects. *Journal of Neurophysiology*, 73, 373-386.

Toni I, Shah NJ, Fink GR, Thoenissen D, Passingham RE, Zilles K. 2002. Multiple movement representations in the human brain: an event-related fMRI study. *Journal of Cognitive Neuroscience*, 14, 769–784.

Tootell RB, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ, Rosen BR, Belliveau JW. 1995. Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience*, 15, 3215-3230.

van Strien, JW. 1992. Classificatie van links- en rechtshandige proefpersonen (Classification of left- and right-handed subjects.). *Nederlands Tijdschrift voor de Psychologie en Haar Grensgebieden* 47, 88-92.

Wang Z, Childress AR, Wang J, Detre JA. 2007. Support vector machine learning-based fMRI data group analysis. *NeuroImage*, 36, 1139-1151.

Zeki, S., & Goodenough, O.R. (2004). Law and the brain: Introduction. *Philosophical Transactions of the Royal Society*, 359, 1661-1665.

Table

Table 1: Areas showing significant activation in GLM analysis

A) No-Decision > NoGo						
<i>Area</i>	<i>BA</i>	<i>peak coordinates</i> <i>(MNI)</i>	<i>Z-score</i>	<i>Extent</i>	<i>p*</i>	
Paracingulate cortex, Right superior frontal gyrus	32 6 8	4, 46, 4 11, 42, 49 21, 35, 49	4.56	232	<0.001	
Precuneus	31 23	4, -32, 42 4, -46, 25	3.79	28	<0.001	
Right superior frontal gyrus	10	18, 63, 21	3.78	23	<0.001	
Right inferior parietal lobule (angular gyrus)	39	56, -53, 25	3.46	21	<0.001	
B) NoGo > No-Decision						
<i>Area</i>	<i>BA</i>	<i>peak coordinates</i> <i>(MNI)</i>	<i>Z-score</i>	<i>Extent</i>	<i>p*</i>	
Left inferior parietal lobule (supramarginal gyrus), Left superior temporal gyrus	40 22	-56, -46, 28 -60, -49, 18 -63, -46, 11	4.48	58	<0.001	
Left dorsal premotor cortex	6	-56, -7, 46	3.99	15	<0.001	

Figure legends

Figure 1: Schematic drawing of the experimental design.

Figure 2: Activation map resulting from the GLM analysis averaged over 15 subjects (uncorrected $p < 0.001$, $k = 10$ voxels) mapped onto a template brain (colin27). Coloured labels indicate positive T values. A) Contrast of NoGo > No-Decision; displayed are activities in left dorsal premotor cortex (dPMC), left inferior parietal lobule (IPL) extending into superior temporal gyrus. B) Contrast of No-Decision > NoGo, displayed are activities paracingulate cortex (parac) and precuneus (prec) on the medial view, right IPL, right superior frontal gyrus (SFG) on the lateral view.

Figure 3: Map resulting from the second-level analysis on classification spatial discriminance maps (SDMs) averaged over 15 subjects (FWE = 0.05, $k = 10$) mapped onto a template brain (colin27). Displayed are activities in right inferior frontal gyrus (IFG), supplementary motor area (SMA), right motion-selective area (V5/MT), right anterior insula/ operculum (AIC) and precuneus (prec).

Figure 1

Figure 2

Figure 3