



Who is honest and why: Baseline activation in anterior insula predicts inter-individual differences in deceptive behavior



Thomas Baumgartner*, Lorena R.R. Gianotti, Daria Knoch

Social and Affective Neuroscience, Department of Psychology, University of Basel, Birmannsgasse 8, CH-4055 Basel, Switzerland

ARTICLE INFO

Article history:

Received 22 January 2013

Accepted 27 May 2013

Available online xxx

Keywords:

Deception

Dishonesty

Honesty

Individual differences

Neural trait marker

Resting EEG

Source localization

Anterior insula

ABSTRACT

Humans engage in deceptive behavior that negatively affects others. The propensity to deceive is, however, characterized by vast inter-individual heterogeneity that is poorly understood. Attempts to investigate the origins of this heterogeneity have so far mainly relied on subjective measures and have shown little predictive power. Here, we used resting electroencephalography to measure objective and stable individual differences in neural baseline activation in combination with an ecologically valid deception paradigm. Results showed that task-independent baseline activation in the anterior insula, a brain area implicated in mapping internal bodily states and in representing emotional arousal and conscious feelings, predicts individuals' propensity for deceptive behavior. The higher the neural baseline activation in this area is, the lower individuals' propensity to deceive. Moreover, results provide evidence that high baseline activation in the anterior insula is associated with negative affect and dispositional tendencies to avoid aversive emotional situations. These results provide converging neural and psychological evidence that individuals might avoid a deceptive act due to a highly active negative emotional system which would make a deceptive act too stressful and bothersome.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

People lie frequently in everyday social interactions (DePaulo, Kirkendol, Kashy, Wyer, & Epstein, 1996). Although not all of these lies are harmful, people often engage in deceptive behavior that negatively affects others (DePaulo, Ansfield, Kirkendol, & Boden, 2004).

Despite the pervasiveness of deception in human society, considerable inter-individual differences in the propensity to deceive can be observed in laboratory and field studies (Kashy & DePaulo, 1996; Phillips, Meek, & Vendemia, 2011). Attempts to explain this inter-individual variation in deceptive behavior primarily focused on psychological variables such as stable personality traits (e.g. Kashy & DePaulo, 1996; McLeod & Genereux, 2008; Phillips et al., 2011) and personality disorders (e.g. Spidel, Herve, Greaves, & Yuille, 2011). Correlations between personality traits and deception, however, have yielded rather variable and inconsistent results. For example, machiavellian personality traits correlated in some studies with deception (e.g. McLeod & Genereux, 2008), but failed to show this correlation in others (e.g. Phillips et al., 2011). One reason for these mixed results might lie in the subjective nature of the measurements employed (i.e. self-reports of personality

traits and deceptive behavior), which are known to be affected by various biases (e.g. demand characteristics, social desirability). Moreover, the predictive power of these personality traits was rather low; often below 5% of variance could be explained by a single personality trait (e.g. DePaulo, 2004). Thus, the use of more objective individual trait measurements in combination with more ecologically valid deception paradigms might be beneficial in understanding the inter-individual differences in the propensity to deceive.

An ideal trait measurement of this type is task-independent neural baseline activation measured by resting electroencephalography (EEG) because this measurement demonstrates high stability over time and high specificity (i.e. the extent to which an EEG pattern uniquely belongs to a given person). Studies investigating the stability of resting EEG revealed test–retest reliabilities of up to 0.8 over a period of up to 5 years (Cannon et al., 2012; Dunki, Schmid, & Stassen, 2000; Gold, Fachner, & Erkkila, 2013; Napflin, Wildi, & Sarnthein, 2007; Smit, Posthuma, Boomsma, & Geus, 2005; Williams et al., 2005) and studies exploring the specificity revealed recognition rates of up to 99% (Dunki et al., 2000; Napflin et al., 2007). Due to high intra-individual stability and specificity, this measurement provides an ideal neural trait marker to investigate the sources of inter-individual differences in deceptive behavior.

To measure variation in deceptive behavior in an ecological valid situation, we used a paradigm resembling an economic exchange situation between two interaction partners – an investor and a trustee. In this paradigm, individuals were free to deceive or to tell

* Corresponding author. Tel.: +41 61 267 0639.

E-mail address: t.baumgartner@unibas.ch (T. Baumgartner).

the truth and all decisions had real consequences for all involved interaction partners. More precisely, the trustee first has to make a promise decision at the beginning of a series of three subsequent trust game trials, indicating whether he/she *always*, *mostly*, *sometimes*, or *never* plans to be trustworthy. In this context, being trustworthy means returning money so that both players earn the same amount. The investor is always informed about the trustee's promise and can then decide (based on the trustee's promise) whether to trust the trustee and invest money or to not trust him/her and thus to keep the initial endowment of 2 money units (MUs). If the investor trusts the trustee, the experimenter increases the amount the investor sends to the trustee by the factor of five. The trustee can then freely decide to be honest and keep the promise, or he/she may also decide to be deceptive and break the promise and thus violate the investor's trust by not returning money.

Previous studies on the neural underpinnings of deceptive behavior focused on brain activation during the decision-making process rather than examining individuals' task-independent neurophysiological characteristics (for recent reviews see Abe, 2011; Sip, Roepstorff, McGregor, & Frith, 2008). Thus, these studies do not identify neural traits responsible for predisposing subjects to deceive or tell the truth. Nevertheless, the findings of these studies can be used to generate hypotheses about the potential neural sources driving individuals' propensity for deceptive and truthful behavior.

One of the most consistent findings in these studies (e.g. Abe, Suzuki, Mori, Itoh, & Fujii, 2007; Baumgartner, Fischbacher, Feierabend, Lutz, & Fehr, 2009; Sip et al., 2010; Spence, Kaylor-Hughes, Farrow, & Wilkinson, 2008) is the increased activation in areas of the prefrontal cortex (e.g. dorsolateral prefrontal cortex, DLPFC, ventrolateral prefrontal cortex, VLPFC) during the deceptive act. Due to the role of these areas in cognitive control (e.g. Miller & Cohen, 2001), response selection (e.g. Rowe, Toni, Josephs, Frackowiak, & Passingham, 2000), and response inhibition (Aron, Robbins, & Poldrack, 2004), these prefrontal activation patterns have commonly been interpreted as to suggest that the suppression of the honest response and the generation of the deceptive response requires control-related processes (Abe, 2011). In line with this assumption is a recent study of patients with Parkinson's disease, which demonstrated that the reduced tendency to deceive in these patients is associated with decreased metabolic rates in the left DLPFC and right anterior prefrontal cortices (Abe et al., 2009). Moreover, studies in patients with a history of pathological lying showed increased white matter volume in the prefrontal cortex (e.g. Yang et al., 2007). These findings together with previous studies showing that the level of neural baseline activation in the lateral prefrontal cortex positively correlates with control abilities (Gianotti, Figner, Ebstein, & Knoch, 2012; Knoch, Gianotti, Baumgartner, & Fehr, 2010) led us to hypothesize that subjects with higher baseline activation in control-related areas of the prefrontal cortex might show an increased propensity to deceive.

Other regions that have been demonstrated to be activated during the deceptive act (e.g. Abe et al., 2007; Baumgartner et al., 2009; Kozel et al., 2005) comprised areas involved in processing emotions (e.g. anterior insula, Craig, 2009), and areas involved in processing conflict (e.g. anterior cingulate cortex, Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999). It has been argued that the decision conflict and the psychological stress (e.g. guilt) that often accompany the generation and enactment of a lie are encoded in these brain areas (Sip et al., 2008). Thus, it might be that differences in the neural functioning of these brain areas affect individuals' propensity to deceive. We hypothesize that subjects with higher baseline activation in emotion- and conflict-related brain areas might show a reduced propensity to lie because a deceptive act would cause too much stress or bothersome emotions (e.g. guilt).

Taken together, we hypothesize that inter-individual differences in the propensity to deceive might be positively associated with baseline activation levels in control-related areas and/or negatively associated with baseline activation levels in emotion- and conflict-related areas.

2. Materials and methods

2.1. Subjects

We measured neural baseline activation and the propensity for deceptive behavior in 50 healthy individuals (mean age \pm SD = 23.3 \pm 4.9, 19 men and 31 female). One male subject was excluded from analyses based on outlier brain data and regression influence statistics (*Cook's Distance* = 0.164), leaving 49 participants for analyses. All subjects were right-handed and had no history of neurological or psychiatric disorders or alcohol and drug abuse. The study was approved by the local ethics committee. All subjects gave written, informed consent and were informed of their right to discontinue participation at any time. Subjects received 40 Swiss francs (CHF 40; CHF 1 = \$1 U.S.) for participating, in addition to the money earned in the deception paradigm.

2.2. Procedure

The deception paradigm and the EEG recordings took place during different sessions, which were separated by several weeks. The deception paradigm was conducted in our behavioral laboratory with interconnected computer terminals, while the EEG recordings were conducted in our EEG laboratory, where only one subject was measured at one time.

2.3. Deception paradigm

Subjects in the role of a trustee played 9 rounds of the deception paradigm with 9 different, anonymous interaction partners in the role of an investor. We deliberately chose anonymous one-shot interactions in order to exclude reputation effects and strategic spillovers across trials. At the beginning of such an interaction, an investor receives an endowment of 2 money units (MUs), whereas the trustee receives nothing. Then, the trustee has to make a promise decision at the beginning of a series of three subsequent trust game trials, indicating whether he/she *always*, *mostly*, *sometimes*, or *never* plans to be trustworthy and return the money. The investor is then informed about the trustee's promise and he/she can decide to send his/her endowment of two MUs to the trustee, or he/she can decide to keep his/her endowment. If the investor keeps his/her endowment, the round ends and the trustee gets nothing. However, if the investor trusts the trustee and invests money (which occurred in most cases [89%] due to the high promise level that nearly all trustees chose, see results section for details), the experimenter increases the amount sent by a factor of five, so that the trustee receives 10 MUs. He/she then can freely decide whether to keep the promise or to break it. If the trustee decides to keep his/her promise and return the investment, both players earn 5 money units (5 money units = CHF 1, about \$1). If the trustee decides to break the promise and keep the investment, he/she earns 10 money units (=CHF 2, about \$2), whereas the investor receives no money in this round. The software package z-Tree, a program for conducting behavioral exchange experiments was used for presenting screens and for collecting behavioral data.

2.4. Questionnaires

We administered the following trait questionnaires: positive and negative affect schedule (PANAS, Watson, Clark, & Tellegen, 1988), behavioral inhibition and behavioral approach system scales (BIS/BAS scales, Carver & White, 1994), and Barratt impulsiveness scale (Patton, Stanford, & Barratt, 1995). The PANAS is a 20-item questionnaire which measures dispositional tendencies to experience negative affect (e.g. distress, unpleasurable engagement) and positive affect (e.g. pleasurable engagement) in life. The BIS/BAS scales consist of 24 items which assess individual dispositional differences in the sensitivity of two general motivational systems underlying behavior. The behavioral approach system (BAS) is believed to regulate appetitive motives, in which the goal is to move toward something desired. A behavioral avoidance (or inhibition) system (BIS) is said to regulate aversive motives, in which the goal is to move away from something unpleasant. The Barratt impulsiveness scale consists of 30 items which assess individuals' self-control abilities, i.e. how good subjects are able to control impulsive behavior. High values on this scale indicate low control ability.

2.5. EEG recording and processing

Subjects were seated comfortably in a dimly lit, quiet room with intercom connection to the experimenter. Then 64 Ag-AgCl active electrodes were placed following the 10-10 montage covering the entire scalp, as recommended by the International Federation of Clinical Neurophysiology Standards for Digital Recording of Clinical EEG (Nuwer et al., 1998). During the recordings, the signals were

referenced to CMS (common mode sense), while DRL (driven right leg) served as ground. Data were recorded with a sampling rate of 512 Hz (24 bit precision; bandwidth: 0.1–100 Hz). Horizontal and vertical eye movements were recorded with electrodes at the left and right outer canthi and left infraorbital.

After placement of the EEG electrodes, subjects were instructed that EEG recording was to be done during resting with open or closed eyes. The protocol consisted of 20-s eyes open followed by 40-s eyes closed, repeated five times. Only data from the 200-s eyes-closed condition were further analyzed.

In order to eliminate artifacts, independent component analysis (ICA) was applied, and ICA components that clearly accounted for vertical and horizontal eye movements were removed from the EEG without topographic distortion. In addition to the rejection of sweeps where any channel exceeded the amplitude of $\pm 100 \mu\text{V}$, the data were visually inspected to reject remaining artifacts, using a moving, non-overlapping 2-second window. For each subject, channels exhibiting substantial noise were interpolated using a 3D spherical spline interpolation procedure. The EEG data were then recomputed against the average reference. All available artifact-free 2-s EEG epochs (on average, 84.4 ± 15.4 epochs available per subjects) were then subjected to conventional spectral analyses. Power spectra were integrated for the following seven independent frequency bands (Kubicki, Herrmann, Fichte, & Freund, 1979): delta (1.5–6 Hz), theta (6.5–8 Hz), alpha1 (8.5–10 Hz), alpha2 (10.5–12 Hz), beta1 (12.5–18 Hz), beta2 (18.5–21 Hz), and beta3 (21.5–30 Hz).

Standardized low-resolution brain electromagnetic tomography (sLORETA, Pascual-Marqui, Esslen, Kochi, & Lehmann, 2002) was used to estimate the intracerebral electrical sources that generated the scalp-recorded activity in each of the seven frequency bands. sLORETA computes electric neuronal activity as standardized current density (unit: amperes per square meter, A/m^2) without assuming a predefined number of active sources. The sLORETA solution space consisted of 6239 voxels (voxel size: $5 \text{ mm} \times 5 \text{ mm} \times 5 \text{ mm}$) and is restricted to cortical gray matter and the hippocampus, as defined by the digitized Montreal Neurological Institute (MNI) probability atlas. The sLORETA method has received considerable validation from studies combining EEG/MEG source localizations performed in conjunction with other localization methods, such as functional Magnetic Resonance Imaging (fMRI, Mobascher et al., 2009; Olbrich et al., 2009) and Positron Emission Tomography (PET, Laxton et al., 2010). Further, the method has been validated with experimental data for which the true generators are known from invasive, implanted depth electrodes (Zumsteg, Friedman, Wieser, & Wennberg, 2006; Zumsteg, Lozano, Wieser, & Wennberg, 2006) and has been demonstrated to be able to correctly localize deep structures such as the anterior cingulate cortex (e.g. Pizzagalli et al., 2001) and mesial temporal lobes (e.g. Zumsteg, Friedman, et al., 2006).

In order to reduce confounds that have no regional specificity, such as inter-subject variability in total power, a global normalization and log-transformation of the sLORETA images was carried out prior to subsequent statistical analyses.

2.6. Statistical analyses of behavioral and brain data

In order to explore the propensity for deceptive behavior, we first calculated the promise rate and the return rate. The promise rate was created by transforming the three 4-graded promise decisions into percentage values (always = 100%, mostly = 66.66%, sometimes = 33.33%, never = 0%) and averaging them. The return rate was created by calculating the percentage of trials where trustees decided to pay back the money, if the investors trusted them. Based on these two rates, we conducted a two-step cluster analysis to explore the inter-individual variability in deceptive behavior.

The main goal of this study was to examine whether subjects' propensity to deceive can be predicted based on a task-independent, neural trait marker. For that purpose, we further created a deception score by subtracting the return rate from the promise rate. Thus, the higher the deception score, the more dishonest the behavior. We conducted whole brain regression analyses (separately for each frequency band) using this deception score as dependent variable. We controlled in this analysis for the number of trials in which the investor did not invest money, i.e. trials with no decision on the trustee's side. The nonparametric randomization approach (Nichols & Holmes, 2002) was used for estimating empirical probability distributions (number of randomizations used: 5000) and the corresponding corrected (for multiple comparisons) critical probability thresholds.

3. Results

3.1. Behavioral results

To explore subjects' propensity for deceptive behavior, we conducted a two-step cluster analysis using both the promise rate and return rate as dependent variables. As expected, this analysis revealed considerable inter-individual variation in deceptive behavior, indicated by a cluster solution with three separated clusters (see Fig. 1). The first group of subjects included 17 individuals (12 female) who demonstrated honest behavior, i.e. they promised to pay back the money in 97.4% of the trials ($\text{SEM} = 1.18$) and kept

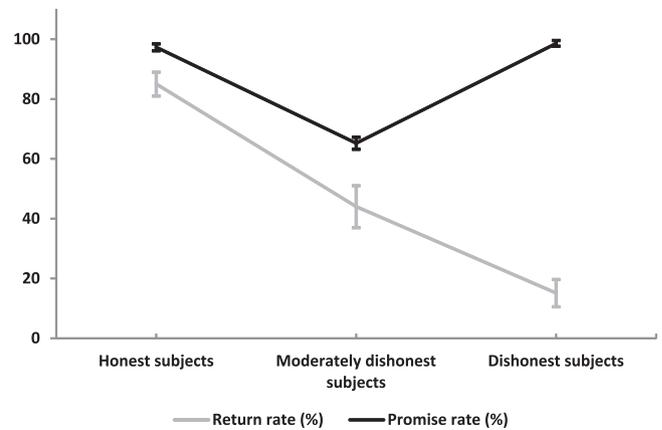


Fig. 1. Behavioral results. Depicted are (A) trustees' average return rate ($\pm \text{SEM}$) and (B) trustees' average promise rate ($\pm \text{SEM}$), broken down by the three types of trustees revealed by the cluster analysis: honest subjects, moderately dishonest subjects, dishonest subjects.

their promise in 85.0% of the trials ($\text{SEM} = 3.98$). In sharp contrast, the second group of subjects included 16 individuals (10 females) who demonstrated dishonest behavior, i.e. they promised to pay back the money in 98.6% of the trials ($\text{SEM} = 0.95$), but only kept their promise in 15.1% of the trials ($\text{SEM} = 4.55$). Thus, they deceived most of the time. Finally, a third group of subjects included 16 individuals (9 females) who demonstrated moderately dishonest behavior, i.e. they promised to pay back the money in 65.3% of the trials ($\text{SEM} = 2.00$) and kept their promise in 44.0% of the trials ($\text{SEM} = 7.04$).

3.2. Brain results

As demonstrated above, subjects strongly differed in their deceptive behavior. Accordingly, we conducted whole brain regression analyses using the deception score as dependent variable. Using sLORETA as a source localization technique to estimate intracerebral sources underlying scalp-recorded resting EEG, we found that in the theta frequency band, 39 voxels showed significant positive correlations between current density and deceptive behavior ($r(46) = .475, p < 0.05$, corrected). As shown in Fig. 2, these voxels all fell into one cluster in the left hemisphere, encompassing mainly the anterior insula, but also extending into the inferior frontal gyrus (MNI coordinates peak voxel: $x = -40, y = 15, z = 0$, Brodmann area 13/47, an area also often labeled as fronto-insular cortex, Craig, 2009). This cluster accounted for 22.6% of the variance in deceptive behavior. Because baseline theta activity is an inverse indicator of neural activation (Oakes et al., 2004), these findings suggest that the higher the baseline activation in the anterior insula, the lower the propensity for dishonest behavior. Moreover, our findings were specific to the left anterior insula because no significant correlations were found in any other EEG frequency bands, and in no other brain region was baseline theta activity correlated with deceptive behavior at the corrected significance threshold. Finally, we tested whether there are gender differences in the correlation between theta activity in the anterior insula and deceptive behavior. For that purpose, we included in our regression model an interaction term of gender \times theta activity along with a main effect of gender, which is necessary to test interaction effects. Findings revealed no evidence that the correlation between theta activity and deceptive behavior is affected by gender (interaction term gender \times theta activity: $p = 0.984$).

In sum, we demonstrated that the level of baseline activation in the anterior insula predicts individuals' propensity to lie. We next tested our hypothesis that increased baseline activation in

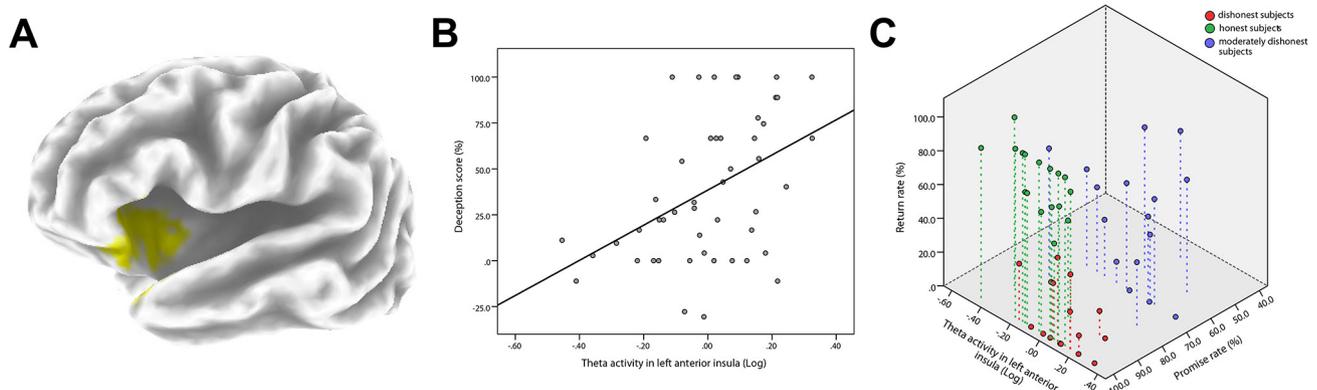


Fig. 2. Baseline theta activity in the left anterior insula predicts deceptive behavior. Depicted in (A) on an inflated brain is the area of the left anterior insula (extending into the inferior frontal gyrus) where baseline theta activity (A/m^2 , log-transformed) is positively correlated with deceptive behavior (at $p < 0.05$, corrected for the whole brain). The scatter plot in (B) illustrates the correlation between baseline, task-independent theta activity in the left anterior insula (based on a 10 mm spherical ROI around the peak of the positive correlation depicted in A) and the deception score (promise rate minus return rate). Because resting theta activity is an inverse indicator of cortical activation, this finding suggests that the higher the cortical activation in the left anterior insula (i.e. the lower the baseline theta activity), the lower the subjects' propensity for deceptive behavior. The scatter plot in (C) further illustrates this finding using each subject's individual return rate (%), promise rate (%) and baseline theta activity in the left anterior insula (log-transformed). Furthermore, this figure shows the three types of trustees derived from the cluster analysis.

this emotion-related brain area might suggest that a deceptive act would become too emotional and stressful for these subjects and thus they might try to avoid such highly aversive emotional situations. Baseline theta activity in the anterior insula was based in these analyses on a spherical ROI (10 mm in diameter) around the peak of the positive correlation depicted in Fig. 2. We found that the trait measure of negative affect correlates negatively with baseline theta activity (NA scale: $r(47) = -0.367$, $p = 0.01$). No correlation was found with the trait measure of positive affect (PA scale: $r(47) = 0.079$, $p = 0.592$), suggesting that the higher the baseline activation in the anterior insula (i.e. power in the theta band is negatively related to activation), the more individuals report to experience distress and unpleasant affect in their life. Moreover, we found that subjects' dispositional tendencies to avoid negative (BIS scale: $r(47) = -0.406$, $p = 0.004$), but not to approach positive (BAS scale: $r(47) = 0.142$, $p = 0.330$) situations in their life is also significant negatively correlated with baseline theta activity in the anterior insula, i.e. the higher the baseline activation in the anterior insula (i.e. power in the theta band is negatively related to activation), the stronger subjects' propensity to avoid negative circumstances in life. Thus, these correlational findings corroborate our assumption that subjects with high baseline activation in the anterior insula might avoid the deceptive act because their emotional system would react too strongly in such an aversive situation. Finally, we found no correlation between baseline theta activity in the anterior insula and subjects' dispositional tendencies to control impulsive behavior ($r(47) = -0.147$, $p = 0.315$).

4. Discussion

Deceiving involves emotional and cognitive processes which lead to strong task-dependent activation in emotion-related and control-related brain areas. Here we demonstrate the first evidence that a specific neural trait marker – task-independent, baseline activation of the anterior insula – is able to predict deceptive behavior in an ecologically valid paradigm with real monetary consequences for both interaction partners. The higher the level of baseline activation in the anterior insula – a brain area implicated in mapping internal bodily states and in representing emotional arousal and conscious feelings (for a recent review Craig, 2009) – the lower individuals' propensity for deceptive behavior. Thus, dishonest subjects, compared to honest subjects, seem to be characterized by a lower level of neural baseline activation in an area that has

been shown to be activated while subjects commit a deceptive act (Baumgartner et al., 2009). Moreover, subjective measurements indicate that the baseline activation in the anterior insula is positively associated with negative affect and dispositional tendencies to avoid aversive emotional situations. Thus, we conjecture that a high level of baseline activation in the anterior insula might predispose individuals to be honest due to a hyperactive emotional system which would make a deceptive act too stressful and bothersome. One psychological component that makes a deceptive act stressful is guilt (Sip et al., 2008). Interestingly, task-dependent activation in the anterior insula has been associated with guilt-aversion (Chang, Smith, Dufwenberg, & Sanfey, 2011). Thus, it might be that high level of baseline activation in the anterior insula indicates increased guilt-aversion, which in turn increases individuals' propensity to be honest.

Notably, an answer to the question of why some people lie more or less than others can only be attempted by using a task-independent measure such as neural baseline activation and a task that does not force a subject to lie. This is because honest individuals' task-dependent activity related to dishonesty cannot be measured without forcing the subjects to deceive. However, forced deception lacks important elements of guilt and psychological stress (Sip et al., 2008) and would therefore not result in a true representation of the honest individuals' activation pattern in the anterior insula.

Moreover, we do not find any evidence in our data that the level of baseline activation in control-related areas of the prefrontal cortex predisposes individuals to deceive. Thus, these findings suggest that the normal functioning of the prefrontal cortex is a critical prerequisite for (successful) lying, but does not affect individuals' propensity to tell a lie or the truth. In line with this interpretation are recent studies demonstrating that disrupting areas of the prefrontal cortex with transcranial direct current stimulation does not affect individuals' propensity to deceive, but changes the speed and efficiency of deceptive responses (Karim et al., 2010; Priori et al., 2008). Moreover, although preschool children can distinguish mistakes from lies and are capable of deception, these children are far from being “perfect liars” (Polak & Harris, 1999; Siegal & Peterson, 1998). In this regard, it is interesting to note that during early childhood when children begin lying, prefrontal cortex development is still in progress (Gogtay et al., 2004). Thus, the immature prefrontal cortex might be the reason that children cannot make use of deception effectively, further supporting our interpretation that

normal functioning of prefrontal cortex is necessary during successful deceptions, but does not predispose individual to lie or tell the truth.

Our findings raise questions ripe for future research: what might have caused these considerable differences in baseline activation in the anterior insula? Because individual differences in EEG power bands are strongly influenced by genetic factors (Zietsch et al., 2007), it is likely that genes contribute to these inter-individual differences. However, it might also be that the environment or a combination of genes and environment drive these differences in neural baseline activation and associated differences in deceptive and honest behavior. Furthermore, it would be interesting to explore whether task-independent insula activation is able to predict individuals' behavioral disposition in other contexts where individuals are required to decide between prosocial and antisocial behavioral options and where task-dependent insula activity has been observed (e.g. Chang et al., 2011; Hein, Silani, Preuschhoff, Batson, & Singer, 2010). Finally, because the LORETA solution space is restricted to the cortical gray matter and the hippocampus, future studies could apply other methods (e.g. structural MRI, resting functional MRI) to explore whether neural trait measures of sub-cortical areas (e.g. amygdala) can also be used to predict deceptive behavior.

In sum, we provide evidence that a specific neural trait marker – task-independent baseline activation in the anterior insula – predicts individuals' propensity for deceptive behavior. Moreover, we provide evidence that high baseline activation in the anterior insula is associated with negative affect and dispositional tendencies to avoid aversive emotional situations. Together these results provide converging psychological and neural evidence as to why some people behave more honest than others. A deceptive act in the applied behavioral paradigm means that subjects deliberately break their promise, abuse the interactions partner's trust and maximize their monetary reward. Previous studies showed that a violation of social norms leads to a negative emotional state (Haidt, 2003). For those people with dispositionally high negative emotional activity and a tendency to inhibit behavior that may lead to undesirable consequences, a lie may be too stressful. They are thus motivated to avoid negative affective states that would be associated with a deceptive act. On a more general note, our results might point to the added value by integrating the different measures of (personality) psychology and neuroscience in the study of the sources of individual differences in complex human behaviors.

Acknowledgements

This project was supported by a grant to DK by the Swiss National Science Foundation (PP00P1_123381). We further acknowledge support from the Forschungsfonds der Universität Basel and the Mens Sana Stiftung.

References

- Abe, N. (2011). How the brain shapes deception: An integrated review of the literature. *Neuroscientist*, 17(5), 560–574.
- Abe, N., Fujii, T., Hirayama, K., Takeda, A., Hosokai, Y., Ishioka, T., et al. (2009). Do parkinsonian patients have trouble telling lies? The neurobiological basis of deceptive behaviour. *Brain: A Journal of Neurology*, 132(Pt 5), 1386–1395.
- Abe, N., Suzuki, M., Mori, E., Itoh, M., & Fujii, T. (2007). Deceiving others: Distinct neural responses of the prefrontal cortex and amygdala in simple fabrication and deception with social interactions. *Journal of Cognitive Neuroscience*, 19(2), 287–295.
- Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, 8(4), 170–177.
- Baumgartner, T., Fischbacher, U., Feierabend, A., Lutz, K., & Fehr, E. (2009). The neural circuitry of a broken promise. *Neuron*, 64(5), 756–770.
- Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, 402(6758), 179–181.
- Cannon, R. L., Baldwin, D. R., Shaw, T. L., Diloreto, D. J., Phillips, S. M., Scruggs, A. M., et al. (2012). Reliability of quantitative EEG (qEEG) measures and LORETA current source density at 30 days. *Neuroscience Letters*, 518(1), 27–31.
- Carver, C. S., & White, T. L. (1994). Behavioral-inhibition, behavioral activation, and affective responses to impending reward and punishment—the BIS/BAS scales. *Journal of Personality and Social Psychology*, 67(2), 319–333.
- Chang, L. J., Smith, A., Dufwenberg, M., & Sanfey, A. G. (2011). Triangulating the neural, psychological, and economic bases of guilt aversion. *Neuron*, 70(3), 560–572.
- Craig, A. D. (2009). How do you feel—now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, 10(1), 59–70.
- DePaulo, B. M. (2004). The many faces of lies. In A. G. Miller (Ed.), *The social psychology of good and evil* (pp. 303–326). New York: Guilford Press.
- DePaulo, B. M., Ansfield, M. E., Kirkendol, S. E., & Boden, J. M. (2004). Serious lies. *Basic and Applied Social Psychology*, 26(2–3), 147–167.
- DePaulo, B. M., Kirkendol, S. E., Kashy, D. A., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*, 70(5), 979–995.
- Dunki, R. M., Schmid, G. B., & Stassen, H. H. (2000). Intraindividual specificity and stability of human EEG: Comparing a linear vs a nonlinear approach. *Methods of Information in Medicine*, 39(1), 78–82.
- Gianotti, L. R., Figner, B., Ebstein, R. P., & Knoch, D. (2012). Why some people discount more than others: Baseline activation in the dorsal PFC mediates the link between COMT genotype and impatient choice. *Frontiers in Neuroscience*, 6, 54.
- Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, 101(21), 8174–8179.
- Gold, C., Fachner, J., & Erkkila, J. (2013). Validity and reliability of electroencephalographic frontal alpha asymmetry and frontal midline theta as biomarkers for depression. *Scandinavian Journal of Psychology*, 54(2), 118–126.
- Haidt, J. (2003). The moral emotions. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook in affective sciences* (pp. 852–870). Oxford: Oxford University Press.
- Hein, G., Silani, G., Preuschhoff, K., Batson, C. D., & Singer, T. (2010). Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron*, 68(1), 149–160.
- Karim, A. A., Schneider, M., Lotze, M., Veit, R., Sauseng, P., Braun, C., et al. (2010). The truth about lying: Inhibition of the anterior prefrontal cortex improves deceptive behavior. *Cerebral Cortex*, 20(1), 205–213.
- Kashy, D. A., & DePaulo, B. M. (1996). Who lies? *Journal of Personality and Social Psychology*, 70(5), 1037–1051.
- Knoch, D., Gianotti, L. R., Baumgartner, T., & Fehr, E. (2010). A neural marker of costly punishment behavior. *Psychological Science*, 21(3), 337–342.
- Kozel, F. A., Johnson, K. A., Mu, Q., Grenesko, E. L., Laken, S. J., & George, M. S. (2005). Detecting deception using functional magnetic resonance imaging. *Biological Psychiatry*, 58(8), 605–613.
- Kubicki, S., Herrmann, W. M., Fichte, K., & Freund, G. (1979). Reflections on the topics: EEG frequency bands and regulation of vigilance. *Pharmakopsychiatrie Neuropsychopharmakologie*, 12(2), 237–245.
- Laxton, A. W., Tang-Wai, D. F., McAndrews, M. P., Zumsteg, D., Wennberg, R., Keren, R., et al. (2010). A phase I trial of deep brain stimulation of memory circuits in Alzheimer's disease. *Annals of Neurology*, 68(4), 521–534.
- McLeod, B. A., & Genereux, R. L. (2008). Predicting the acceptability and likelihood of lying: The interaction of personality with type of lie. *Personality and Individual Differences*, 45(7), 591–596.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167–202.
- Mobascher, A., Brinkmeyer, J., Warbrick, T., Musso, F., Wittsack, H. J., Stoermer, R., et al. (2009). Fluctuations in electrodermal activity reveal variations in single trial brain responses to painful laser stimuli—A fMRI/EEG study. *Neuroimage*, 44(3), 1081–1092.
- Napflin, M., Wildi, M., & Sarnthein, J. (2007). Test-retest reliability of resting EEG spectra validates a statistical signature of persons. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 118(11), 2519–2524.
- Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. [Research Support, Non-U.S. Gov't]. *Human Brain Mapping*, 15(1), 1–25.
- Nuwer, M. R., Comi, G., Emerson, R., Fuglsang-Frederiksen, A., Guerit, J. M., Hinrichs, H., et al. (1998). IFCN standards for digital recording of clinical EEG. International Federation of Clinical Neurophysiology. *Electroencephalography and Clinical Neurophysiology*, 106(3), 259–261.
- Oakes, T. R., Pizzagalli, D. A., Hendrick, A. M., Horras, K. A., Larson, C. L., Abercrombie, H. C., et al. (2004). Functional coupling of simultaneous electrical and metabolic activity in the human brain. *Human Brain Mapping*, 21(4), 257–270.
- Olbrich, S., Mulert, C., Karch, S., Trenner, M., Leicht, G., Pogarell, O., et al. (2009). EEG-vigilance and BOLD effect during simultaneous EEG/fMRI measurement. *Neuroimage*, 45(2), 319–332.
- Pascual-Marqui, R. D., Esslen, M., Kochi, K., & Lehmann, D. (2002). Functional imaging with low-resolution brain electromagnetic tomography (LORETA): A review. *Methods & Findings in Experimental & Clinical Pharmacology*, 24 (Suppl. C), 91–95.
- Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology*, 51(6), 768–774.
- Phillips, M. C., Meek, S. W., & Vendemia, J. M. C. (2011). Understanding the underlying structure of deceptive behaviors. *Personality and Individual Differences*, 50(6), 783–789.

- Pizzagalli, D., Pascual-Marqui, R. D., Nitschke, J. B., Oakes, T. R., Larson, C. L., Abercrombie, H. C., et al. (2001). Anterior cingulate activity as a predictor of degree of treatment response in major depression: Evidence from brain electrical tomography analysis. *The American Journal of Psychiatry*, *158*(3), 405–415.
- Polak, A., & Harris, P. L. (1999). Deception by young children following noncompliance. *Developmental Psychology*, *35*(2), 561–568.
- Priori, A., Mameli, F., Cogiamanian, F., Marceglia, S., Tiriticco, M., Mrakic-Sposta, S., et al. (2008). Lie-specific involvement of dorsolateral prefrontal cortex in deception. *Cerebral Cortex*, *18*(2), 451–455.
- Rowe, J. B., Toni, I., Josephs, O., Frackowiak, R. S., & Passingham, R. E. (2000). The prefrontal cortex: Response selection or maintenance within working memory? *Science*, *288*(5471), 1656–1660.
- Siegal, M., & Peterson, C. C. (1998). Preschoolers' understanding of lies and innocent and negligent mistakes. *Developmental Psychology*, *34*(2), 332–341.
- Sip, K. E., Lyngø, M., Wallentin, M., McGregor, W. B., Frith, C. D., & Roepstorff, A. (2010). The production and detection of deception in an interactive game. *Neuropsychologia*, *48*(12), 3619–3626.
- Sip, K. E., Roepstorff, A., McGregor, W., & Frith, C. D. (2008). Detecting deception: The scope and limits. *Trends in Cognitive Sciences*, *12*(2), 48–53.
- Smit, D. J., Posthuma, D., Boomsma, D. I., & Geus, E. J. (2005). Heritability of background EEG across the power spectrum. *Psychophysiology*, *42*(6), 691–697.
- Spence, S. A., Kaylor-Hughes, C., Farrow, T. F., & Wilkinson, I. D. (2008). Speaking of secrets and lies: The contribution of ventrolateral prefrontal cortex to vocal deception. *Neuroimage*, *40*(3), 1411–1418.
- Spidel, A., Herve, H., Greaves, C., & Yuille, J. C. (2011). 'Wasn't me!' A field study of the relationship between deceptive motivations and psychopathic traits in young offenders. *Legal and Criminological Psychology*, *16*(2), 335–347.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect—The Panas Scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070.
- Williams, L. M., Simms, E., Clark, C. R., Paul, R. H., Rowe, D., & Gordon, E. (2005). The test-retest reliability of a standardized neurocognitive and neurophysiological test battery: Neuromarker. *The International Journal of Neuroscience*, *115*(12), 1605–1630.
- Yang, Y., Raine, A., Narr, K. L., Lencz, T., LaCasse, L., Colletti, P., et al. (2007). Localisation of increased prefrontal white matter in pathological liars. *British Journal of Psychiatry*, *190*, 174–175.
- Zietsch, B. P., Hansen, J. L., Hansel, N. K., Geffen, G. M., Martin, N. G., & Wright, M. J. (2007). Common and specific genetic influences on EEG power bands delta, theta, alpha, and beta. *Biological Psychology*, *75*(2), 154–164.
- Zumsteg, D., Friedman, A., Wieser, H. G., & Wennberg, R. A. (2006). Propagation of interictal discharges in temporal lobe epilepsy: Correlation of spatiotemporal mapping with intracranial foramen ovale electrode recordings. *Clinical Neurophysiology*, *117*(12), 2615–2626.
- Zumsteg, D., Lozano, A. M., Wieser, H. G., & Wennberg, R. A. (2006). Cortical activation with deep brain stimulation of the anterior thalamus for epilepsy. *Clinical Neurophysiology*, *117*(1), 192–207.