NeuroImage 81 (2013) 317–324

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

# Impartiality in humans is predicted by brain structure of dorsomedial prefrontal cortex

Thomas Baumgartner *, Bastian Schiller, Christopher Hill, Daria Knoch *

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

ARTICLE INFO

*Article history:*
Accepted 7 May 2013
Available online 18 May 2013

*Keywords:*
Impartiality
Partiality
Outgroup hostility
Ingroup favoritism
Justice
Third-party punishment
Brain anatomy
Social cognition
Dorsomedial prefrontal cortex
Voxel-based morphometry (VBM)
Freesurfer

ABSTRACT

The moral force of impartiality (i.e. the equal treatment of all human beings) is imperative for providing justice and fairness. Yet, in reality many people become partial during intergroup interactions; they demonstrate a preferential treatment of ingroup members and a discriminatory treatment of outgroup members. Some people, however, do not show this intergroup bias. The underlying sources of these inter-individual differences are poorly understood. Here we demonstrate that the larger the gray matter volume and thickness of the dorsomedial prefrontal cortex (DMPFC), the more individuals in the role of an uninvolved third-party impartially punish outgroup and ingroup perpetrators. Moreover, we show evidence for a possible mechanism that explains the impact of DMPFC's gray matter volume on impartiality, namely perspective-taking. Large gray matter volume of DMPFC seems to facilitate equal perspective-taking of all sides, which in turn leads to impartial behavior. This is the first evidence demonstrating that brain structure of the DMPFC constitutes an important source underlying an individual's propensity for impartiality.



## Introduction

Justice's most iconic figure is *lady justice*, an allegorical representation of the moral forces lying at the heart of the ideal provider of justice. *Lady justice* is often depicted as blindfolded to shield her judgment from morally irrelevant information, such as group affiliation. Yet, when granted the position of provider of justice if the own group is implicated in a conflict, ignoring the blindfold and falling into partiality is human nature (Bowles, 2009; Brewer, 1999; Haushofer et al., 2010). This unequal treatment of ingroup and outgroup members, i.e. ingroup favoritism and outgroup hostility, has been documented in many laboratory and field studies in psychology, sociology, and economics (e.g. Bernhard et al., 2006; Brewer, 1979; Halevy et al., 2008; Levine et al., 2005; Tajfel et al., 1971). Despite the widespread occurrence of this intergroup bias, however, there is considerable inter-individual heterogeneity in the degree of this bias. This raises the question: What sets apart impartial people (i.e. people who treat ingroup and outgroup members equally) from those whose judgments are biased in favor of their ingroup?

There is a long psychological tradition of relating personality to differences in partiality (e.g. Batson and Burris, 1994; Graham et al., 2011; Hewstone et al., 2002; Kreindler, 2005; Pratto and Shih, 2000). However, correlations between these personality difference measures and partiality are rather low, suggesting that personality measures generally have limited predictive power (Hewstone et al., 2002). The use of more objective individual markers might therefore help explain inter-individual differences in the propensity for impartiality. Recent applications of brain morphometry indicate that individual differences in brain structure might be such a useful, objective marker because brain structure has been demonstrated to be relatively stable over time in healthy adults and can be used to predict individual differences in various traits (e.g. Baur et al., 2012; DeYoung et al., 2010), skills (e.g. Jancke et al., 2009; Steinbeis et al., 2012), and behavior (e.g. Bickart et al., 2011; Ersche et al., 2012; Morishima et al., 2012). No previous study, however, has examined whether variables reflecting neuroanatomical individual differences, such as gray matter volume or cortical thickness, may help predict individual differences in human's propensity for impartiality.

Previous studies on the neural underpinnings of partiality measured brain activity during the decision-making process rather than examining task-independent neuroanatomical characteristics. Thus, it is difficult to derive clear hypotheses based on these studies. Nevertheless, these studies do allow for speculation about the potential neural structures driving the propensity for impartiality. These studies (e.g. Baumgartner et al., 2012; Falk et al., 2012; Harris and Fiske, 2006) showed that differences in judgment of and behavior towards ingroup and outgroup members are associated with differential activity patterns in areas known to play a key role in social cognition

* Corresponding authors.
*E-mail addresses:* t.baumgartner@unibas.ch (T. Baumgartner), daria.knoch@unibas.ch (D. Knoch).

1053-8119/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2013.05.047

(Adolphs, 2003; Van Overwalle, 2009), including the dorsomedial prefrontal cortex (DMPFC) and bilateral temporo-parietal junction (TPJ).

In order to investigate whether anatomical differences in certain brain structures explain inter-individual differences in impartiality, we applied structural magnetic resonance imaging and measured a person's propensity for impartiality with a third-party punishment paradigm. In this paradigm, judges in the role of an uninvolved third-party were confronted with norm-abiding and norm-violating behavior committed by both ingroup and outgroup members of real social groups (see Material and methods section for details) and had to decide whether to punish this behavior at their own expense (see Fig. 1).

More precisely, subjects in the role of a third-party (player C) were given the opportunity to punish the behavior of players who had previously played a prisoner's dilemma game (PDG). In the PDG, players A and B (either ingroup members, or outgroup members) were each endowed with 20 points and each had to decide simultaneously whether to keep all of the points or to pass them to the other player. Passed points were doubled. Thus, keeping the points equals defection (denoted as D) and passing the points equals cooperation (denoted as C). For example, if player A retained the 20 points while player B transferred the 20 points (behavioral pattern DC), player A earned a total of 60 points (40 points from the transfer plus the initial endowment of 20 points) and player B earned nothing. In order to be able to punish the decisions made by players A and B in the PDG, subjects in the role of player C received an endowment of 10 points at the beginning of each punishment trial. Assigning 1 punishment point cost player C 1 point and cost the sanctioned player 3 points. Points not used for punishment could be retained as income. Notably, we only allowed player C to punish the behavior of one player (either A or B) during each of the punishment trials played. In order to simplify the nomenclature, we recoded all of player C's decisions such that player A always refers to the player that C can punish, while player B always refers to the player that C cannot punish.

To measure subjects' propensity for impartiality, player C was confronted with two different group situations (depicted in Fig. 1). In the group situation OUT/IN, player A was an outgroup member and player B was an ingroup member, whereas in the group situation IN/OUT, player A was an ingroup member and player B was an outgroup member. Comparing player C's punishment decisions between these two group situations reveals player C's propensity for impartiality. Thus, we calculated a partiality score by subtracting punishment points in IN/OUT from punishment points in OUT/IN, separately for all possible behavioral decisions made by players A and B in the PDG (CC, CD, DC, DD). High values on this partiality score indicate that the third-party judges (player C) strongly differed in the treatment of ingroup and outgroup members, i.e. they showed a pronounced tendency towards partiality. Low values on this score indicate that the third-party judges treated ingroup and outgroup members equally, i.e. they demonstrated an impartial punishment pattern. We used this partiality score in order to examine whether inter-individual differences in the propensity for impartiality can be predicted by differences in brain anatomy.

## Materials and methods

### *Subjects*

56 healthy subjects were studied (mean age ± S.D. = 22.3 ± 3.47 years, 26 females, 30 males). Subjects gave informed written consent prior to participating in the study, which was approved by the local ethics committee. No subject had a history of psychiatric illness or neurological disorders. Subjects received 40 Swiss Francs (CHF 40; CHF 1 = about $1 U.S.) for participating, in addition to the money earned in the third-party punishment paradigm.

### *Social groups and ingroup identification scale*

We decided to use naturally occurring social groups. We recruited strong supporters of either soccer clubs (n = 16) or political parties (n = 40) because previous studies using these groups have reported strong behavioral intergroup biases (Ben-Ner et al., 2009; Hein et al., 2010; Koopmans and Rebers, 2009). Subjects in the role of an uninvolved third-party (player C) were given the opportunity to punish supporters of their own or a corresponding rival social group. Note that soccer supporters always interacted with other soccer supporters and political supporters always interacted with other political supporters. Independent t-tests revealed that the two social groups did not differ

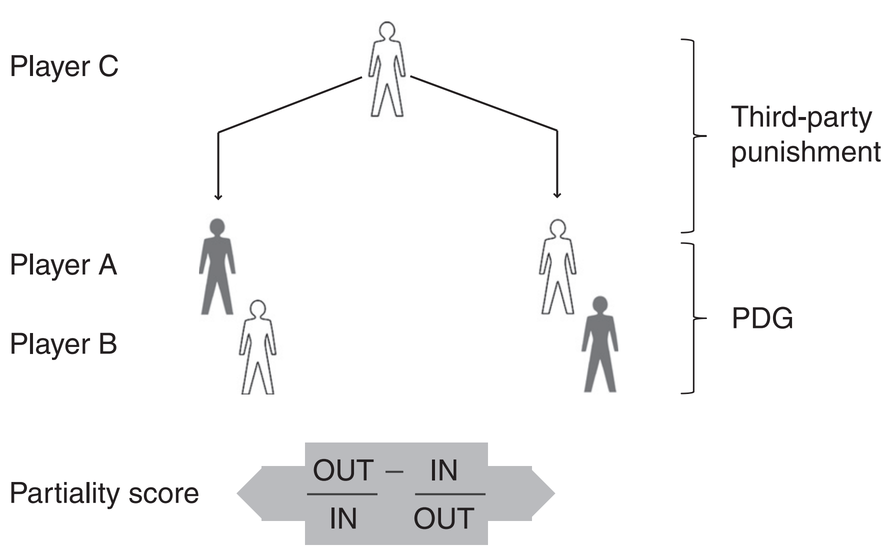


**Fig. 1.** Schematic representation of the study design. Depicted is the applied third-party punishment paradigm. Player C in the role of a third-party judge was confronted with decisions of player A and player B in a Prisoner's Dilemma Game (PDG) and had the opportunity to assign punishment points to player A. Players A and B either stemmed from the same social group as the third-party judge (ingroup members depicted in white colors) or from a different social group (outgroup members depicted in gray colors). In total, third-party judges were confronted with two different group situations: player A is an outgroup member and player B is an ingroup member (termed OUT/IN) or player A is an ingroup member and player B is an outgroup member (termed IN/OUT). Comparing third-party judges' punishment decisions in these two group situations (OUT/IN–IN/OUT) reveals their propensity for impartiality, quantified in the partiality score: high values indicate strong tendencies to partiality and low values indicate strong tendencies to impartiality.