Psychiatry Research: Neuroimaging 174 (2009) 40–46

Contents lists available at ScienceDirect

Psychiatry Research: Neuroimaging

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

Test–retest reliability of event-related functional MRI in a probabilistic reversal learning task☆

Tobias Freyer[a,*], Gabriele Valerius[b], Anne-Katrin Kuelz[a], Oliver Speck[c], Volkmar Glauche[d], Michael Hull[a], Ulrich Voderholzer[a]

[a]Department of Psychiatry, University Medical Centre, Albert-Ludwigs-University Freiburg, Hauptstraße 5, 79104 Freiburg, Germany
[b]Department of Psychosomatic Medicine, Central Institute for Mental Health, 68159 Mannheim, Germany
[c]Department of Biomedical Magnetic Resonance, Institute for Experimental Physics, Otto-von-Guericke-University Magdeburg, Leipziger Str. 44, 39120 Magdeburg, Germany
[d]Department of Neurology, University Medical Centre, Albert-Ludwigs-University Freiburg, Hugstetter Str. 55, 79104 Freiburg, Germany

ARTICLE INFO

Article history:
Received 11 September 2007
Received in revised form 28 November 2008
Accepted 8 March 2009

Keywords:
Functional magnetic resonance imaging
Longitudinal study
Intra-class correlation

ABSTRACT

Repeated functional magnetic resonance imaging (fMRI) studies aim to detect changes in brain activity over time, e.g. to analyze the cerebral correlates of therapeutic interventions. This approach requires a high test–retest reliability of the measures used to rule out incidental findings. However, reliability studies, especially for cognitive tasks, are still difficult to find in the literature. In this study, 10 healthy adult subjects were scanned in two sessions, 16 weeks apart, while performing a probabilistic reversal learning task known to activate orbitofrontal–striatal circuitry. We quantified the reliability of brain activation by computing intra-class correlation coefficients. Group analysis revealed a high concordance for activation patterns in both measurements. Intra-class correlation coefficients (ICCs) were high for brain activation in the associated regions (dorsolateral prefrontal, anterior prefrontal/insular and cingulate cortices), often exceeding 0.8. We conclude that the probabilistic reversal learning task has a high test–retest reliability, making it suitable as a tool for evaluating the dynamics of deterioration in orbitofrontal–striatal circuitry, e.g. to illustrate the course of a psychiatric disorder.

1. Introduction

During the last decade functional magnetic resonance imaging (fMRI) has been established as a popular tool for non-invasive examination of the working human brain. Today a growing number of longitudinal fMRI experiments are designed, e.g. to analyze the progress of a neuropsychiatric disorder, the functional re-organization of the brain after apoplectic stroke or the cerebral correlates of therapeutic interventions. All of these approaches postulate that fMRI constitutes a valid and reliable method, in the sense that differences between measurements at different time points are solely effects of interest and not random or systematic effects produced by the demanding method itself. Considering the large number of publications, there are still few reports about retest reliability measures of fMRI experiments.

Prior studies have used various approaches for evaluating the reproducibility of fMRI signals, differing primarily in the examined brain function: Experiments range from sensorimotor control (Yetkin et al., 1996; Loubinoux et al., 2001; Maitra et al., 2002; Yoo et al., 2005), visual stimulation (Rombouts and Barkhof, 1997; Miki et al., 2000), fear and disgust processing (Stark et al., 2004), auditory odd-ball processing (Kiehl and Liddle, 2003), language production (Brannen et al., 2001; Rutten et al., 2002), verbal (Manoach et al., 2001; Wei et al., 2004; Wagner et al., 2005) and spatial (Casey et al., 1998) working memory tasks to different higher cognitive tasks (McGonigle et al., 2000; Aron et al., 2006). Furthermore, these studies vary broadly in the test–retest interval (from a few hours to more than 1 year) and in the mathematical approach used to determine reproducibility. Several studies qualitatively assessed the consistency of suprathreshold activations in predefined brain areas only and showed mostly analogue results over repeated measurements. For quantitative analyses, many different measures were evaluated in order to determine the reliability: e.g., number of activated voxels, overlap ratio, correlation of activation values or lateralisations, intra-class correlation coefficient (ICC), intersect maps and conjunction analysis. Recently, the computation of ICCs, which index the degree of correlation between subjects at different time points by relating between-subject and total variance, has been proposed as the most exact approach to assess within-subject variability (Manoach et al., 2001; Aron et al., 2006). Therefore, we calculated ICCs of signal changes in previously determined regions of interest, which derived from activations at the group level for either session 1 or session 2 (inclusively).

To our knowledge there are only a few studies quantitatively examining the test–retest reliability of fMRI procedures using higher cognitive tasks (McGonigle et al., 2000; Aron et al., 2006). The present study

☆ Location of work: Department of Diagnostic Radiology, Medical Physics, University Hospital, Hugstetter Str. 55, D-79106 Freiburg, Germany.

* Corresponding author. Tel.: +49 761 270 5040; fax: +49 761 270.
E-mail address: tobias.freyer@uniklinik-freiburg.de (T. Freyer).

0925-4927/$ – see front matter © 2009 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.pscychresns.2009.03.003

aimed to establish the test–retest reliability of fMRI in a probabilistic reversal learning task, which requires subjects to adapt their response strategy according to changes in stimulus–reward contingencies. These set-shifting abilities are of interest in exploring psychiatric disorders, e.g. obsessive–compulsive disorder (OCD), which has been shown to be associated with executive dysfunctions including set-shifting disabilities (Kuelz et al., 2004). Interestingly, neuroimaging studies of OCD demonstrate alterations in orbitofrontal cortex, prefrontal cortices, anterior cingulate cortex and the basal ganglia (Pujol et al., 2004; Mitterschiffthaler et al., 2006), structures that have been shown to be involved in probabilistic reversal learning (Cools et al., 2002; Remijnse et al., 2005). In a first behavioral experiment we found prolonged reaction times with increasing severity of compulsions in OCD patients (Valerius et al., 2008). Remijnse et al. were the first to conduct an fMRI experiment with the reversal learning task comparing OCD patients with healthy subjects, and they found behavioral impairments as well as reduced activation of the left posterior orbitofrontal cortex (OFC), bilateral anterior prefrontal cortex (PFC), bilateral dorsolateral prefrontal cortex (DLPFC) and bilateral insula in patients (Remijnse et al., 2006). Assuming that these results are reproducible and the method is reliable, it should be interesting to examine patients in the course of their disease or before and after cognitive-behavioral psychotherapy.

In this longitudinal event-related fMRI study, we examined the test–retest reliability of a probabilistic reversal learning task, hypothesizing that this task shows minor practice effects and produces stable activation patterns in prefrontal, insular, cingulate and striatal cortices, making it suitable as a tool for evaluating the dynamics of dysfunctional fronto-striatal brain activity due to psychiatric disorders.

2. Methods

2.1. Subjects

Ten right-handed subjects (4 female, $M_{age}=39.8$ years, S.D. = 10.03 years) with no history of psychiatric or neurologic disorder and with normal or corrected-to-normal vision participated. Subjects did not take any psychotropic medication, and there was no substance abuse in the medical history. Informed written consent was obtained from each subject after the procedure had been fully explained. The study was approved by the Ethics Committee of the University of Freiburg according to the guidelines of the Declaration of Helsinki and was conducted in the Department of Diagnostic Radiology at the University Hospital Freiburg.

The subjects performed a Reversal Learning Task twice on two examination days, 16 weeks apart, which is the mean duration of inpatient therapy for patients with OCD. Hardware setup, stimuli presentation, examination time and performing physician (T.F.) were identical in both scanning sessions to control for possible confounders.

2.2. Procedures

2.2.1. Stimuli and task design

Details of the behavioral task have been previously published (Valerius et al., 2008). In brief, subjects were asked to choose one of two simultaneously presented abstract patterns (square and triangle) on each trial by pressing either a left or a right button on a button box positioned on the abdomen of the subject. A feedback in the form of a green happy face or a sad red face was presented immediately after the choice, indicating whether the answer was correct or incorrect. After 10 to 15 (randomised) correct responses, the target object changed and subjects had to adapt their strategy by responding to the formerly incorrect stimulus. To distract subjects, probabilistic errors were interspersed, indicating a wrong choice despite of a correct response. If subjects changed their strategy after a probabilistic error, it was counted as a mistake (SCAPE = strategy change after a probabilistic error). Subjects were asked to avoid mistakes, but still respond as fast as possible. Two successive 9-min sessions with 10 discrimination phases (and therefore 9 reversal stages) each were presented. Each discrimination phase contained between 0 and 4 probabilistic errors. Objects appeared for 2 s during which time subjects had to respond. Feedback was presented for 0.5 s, followed by a fixation cross. The inter-stimulus-interval was 3.3 s (Fig. 1).

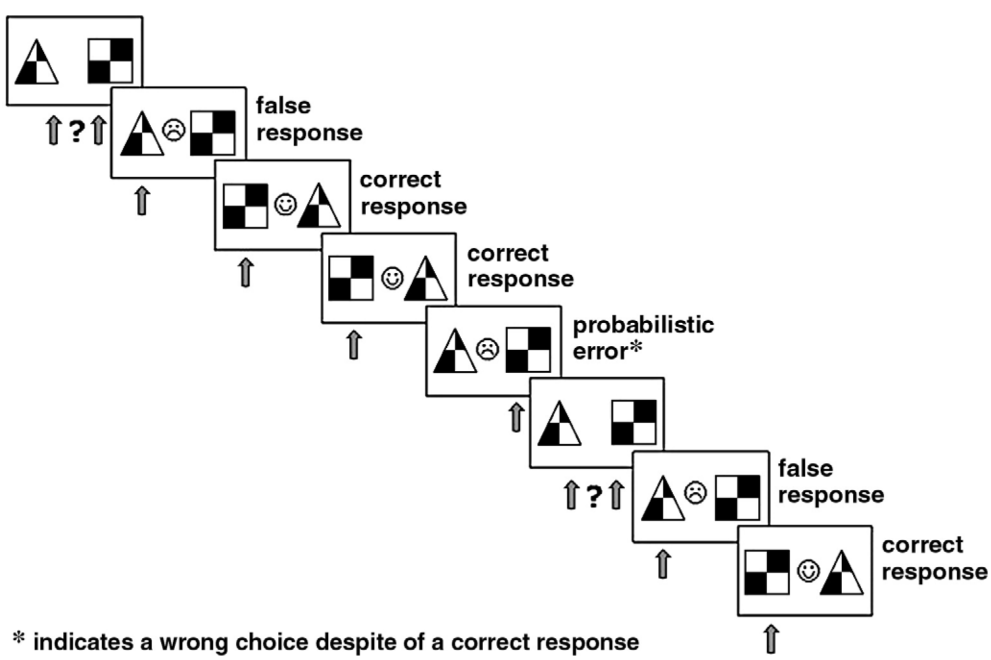

Fig. 1. The reversal learning task: a square and a triangle are presented on a computer screen. Subjects have to respond to one of these objects by using either a left or a right button-press depending on whether their chosen object is on the left or right side of the screen. After their response, feedback in the shape of either a green smiling face or a red sad face appears, indicating whether their reaction was correct or wrong. After 10 to 15 correct responses the strategy changes and subjects have to adapt their reactions and respond to the formerly wrong stimulus. To distract subjects, probabilistic errors are interspersed, indicating a wrong choice in spite of a correct response.