Chapter 20
Functional Magnetic Resonance Imaging (fMRI)

Sebastian Markett

Abstract Functional magnetic resonance imaging (fMRI) is undoubtedly one of the most common techniques used in the cognitive neurosciences and neuroeconomics. The methods section of fMRI papers are oftentimes filled with jargon. We hope to clarify this jargon by defining and explaining the most fundamental concepts. The present chapter has been written to target a broad audience of scholars and students and explains the principles of fMRI: The reader will learn what signals are measured in fMRI, how this measure relates to neural activity, and how fMRI data are most commonly analyzed. This includes a brief summary of physical, physiological, and statistical ideas. We further present a comprehensive step by step guide through a typical fMRI data analysis to provide scholars and students with the appropriate knowledge to understand basic fMRI methodology in research papers and to judge whether the presented analysis is meaningful and appropriately protected against the most common pitfalls in the field of neuroimaging.

20.1 Introduction

Functional magnetic resonance imaging (fMRI) has been the major backbone of the cognitive neurosciences since their very early days. Therefore, it is of little wonder that this method has become extremely popular in the field of neuroeconomics as well. A search with the keywords "functional magnetic resonance imaging" and "neuroeconomics" carried out in Google scholar in early 2016 returned 3120 hits, approximately a quarter of the hits of a search for "neuroeconomics" on its own. What has made this research approach so popular? There are four certain reasons: (a) fMRI has an excellent spatial resolution that allows for the precise anatomical location of neural activation within the brain (b) fMRI comes with sufficient temporal resolution to detect neural correlates of behavior on the basis of experimental trials (c) fMRI is very sensitive and can therefore measure subtle differences in

S. Markett (✉)
Department of Psychology, University of Bonn, Bonn, Germany
e-mail: sebastian.markett@uni-bonn-diff.de

neural activation between experimental conditions which is a prerequisite to test theories on human behavior, and (d) fMRI is noninvasive and safe to use in human research participants because it does not require any pharmacological contrast agents or the lowering of signal detecting devices into the cranium.

There is probably a fifth reason less persuasive to the critical scholar but with great impact for the presentation of research results: fMRI outputs beautiful and intuitively comprehendible images. Even though there are many different ways to present fMRI data, the most common approach to visualize the results is to mark activated regions with red- and yellow-colored blobs on an otherwise greyscale brain. It is these images that has led to the popular notion that fMRI enables the researcher to observe the living brain in action. This might be true to some extent but carries one misconception: The colored blobs themselves are no physiological signals returned by the MRI scanner.[1] They stand at the end of many time-consuming processing and statistical analysis steps and are nothing more than statistical parameters that reflect differences in signal strength between experimental conditions. It is only after the analysis that these statistical parameters are color-coded and then spatially overlaid on a three-dimensional image of the brain. This explains why this approach to neurophysiological data has been labeled statistical parametric mapping and the resulting images statistical parametric maps (Friston et al. 1995).

This chapter is organized in two parts. The first part will focus on the fundamentals of fMRI to answer the question what signal fMRI scanners actually measure and how this signal relates to psychological processes. Understanding the fMRI signal requires some basic knowledge of physics and cell physiology which we hope to cover up in a comprehensive way for readers who have a background in behavioral economics and psychology. The second part will focus on data analysis and will deal with the processing pathway from the raw fMRI data that come out of the MRI scanner to the well-known statistical parametric maps mentioned earlier. Understanding the analysis of fMRI data requires some basic knowledge in psychological experimental design and statistics which we hope to cover up in a comprehensive way for readers without a background in behavioral economics or psychology.

20.2 Fundamentals of fMRI

fMRI is a four-letter acronym. In the introduction, we have already dealt with the fourth letter, the i for imaging, and established that fMRI outputs images of the brain. As a fact, fMRI outputs functional images of the brain. This is what the f stands for and it means that the images acquired in an fMRI scanner allow for inferences on brain function, in this case on neural activity. The opposite (or better

---

[1] fMRI is a special case of magnetic resonance imaging. Functional (fMRI) and structral (MRI) images are acquired on the same scanning device.

the compliment) to fMRI would be structural MRI, an approach not sensitive for brain function but for brain anatomy (see the chapter by Christian Gaser in this edition for the role of structural MRI in the context of neuroeconomics). The remaining two letters in fMRI, the m and the r, stand for magnetic and resonance respectively. They refer to the means by which an fMRI scanner acquires the images: fMRI scanners measure magnetic properties of atomic nuclei in the brain which they accomplish by applying magnetic fields oscillating at the resonant frequency of these nuclei. We will come back to this later in more detail.

The main question we seek to answer in this part of the chapter is how fMRI scanners measure neural activity. As a matter of fact, we can answer this question quite easily on the spot: They do not. This information might come surprising because we usually speak of neural activity revealed by fMRI but it is true: fMRI scanners do not measure neural activity directly. What they do measure, however, are magnetic properties of brain tissues that depend on physiological processes that are most strongly correlated with the neural activity underlying psychological processes.

20.2.1   The Magnet

All fMR imaging starts with a magnet. We have established earlier that fMRI relies on the measurement of magnetic properties of atomic nuclei in the brain. This may sound odd at first glance: If you ever tried to attach a magnet (like the ones that people use to stick notes to their fridge) to your head it will probably come off instantly. This is because the head and the brain have no magnetic properties by themselves. What has magnetic properties, however, are the nuclei of atoms in the brain. MR image acquisition is based on the fact that some atomic nuclei spin around themselves. Hydrogen—the most abundant atom in the brain—has such a spinning nucleus and can therefore be measured by MRI. The nucleus of a hydrogen atom consists of only one positively charged proton. Because of its positive charge, the proton creates a tiny magnetic moment when it spins around itself. This magnetic moment points in the same direction as the proton's spin axis. Under normal conditions, the protons' spin axes will point in random directions, which mean that the same will apply for the magnetic moments. MR imaging does not measure the magnetic moments of single nuclei but the sum of all magnetic moments which is called the net magnetization. Thus, if we tried to measure the magnetic moments under these normal conditions, we would not be able to pick up any signal because the moments would cancel each other out. This is where the magnet enters the stage: If we put our sample (with the containing hydrogen atoms) into a magnetic field, the spins will start to revolve around an axis that is parallel to the magnetic field. This additional spin is called precession spin. You can think about a nucleus' behavior in the magnetic field and the two spins (the regular and the precession spin) as a spinning top (like the ones you may have used to play with as a kid). A spinning top does not only spin around its own axis, it also precesses around a second axis parallel to the earth's gravitational field. If you would watch

Fig. 20.1 Spin of a hydrogen nucleus around its own axis (**a**). When a magnetic field is applied (**b**), the nucleus falls into an additional precession spin in the transversal plane (**c**)
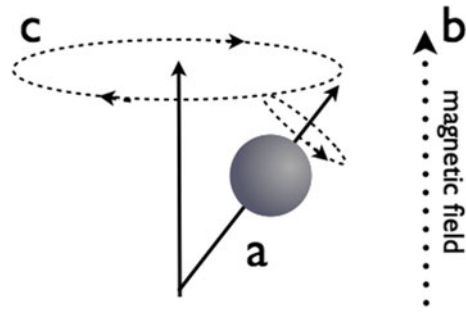
the spinning top from above you would see that the precession spin traces a circle perpendicular to the gravitational field. The precession spins of the nuclei behave in a similar way, only that they do not align with the earth's gravitational field but with the magnetic field applied by the MRI scanner. The axis around which the nuclei process is called the longitudinal direction and the plane in which they precess is called the transversal plane (see Fig. 20.1).

The precession axes align with the magnetic field in two different ways: Either parallel or antiparallel to the magnetic field. The two states differ regarding their energy levels: The parallel state is a low-energy state and is therefore the preferred state of the nuclei. Nonetheless, at each point of time, many nuclei will also spin in the high-energy antiparallel state. Every now and then each nucleus will change its state and flip from the parallel to the antiparallel spin and vice versa.

The more nuclei spin in the parallel relative to the antiparallel state, the higher is the net magnetization in the sample. To get MRI to work we therefore need all (or most) of the nuclei in the parallel state. This can be accomplished by two means. The first approach would be to cool down the sample to the point where no or only little molecular motion occurs. This, however, would be way too cold for the living brain and is therefore not practical for our purpose. The other approach is the one used in MRI scanners: If we dramatically increase the field strength of our magnet, the vast majority of nuclei will align their precession spins with the magnetic field in parallel. The field strength of strong magnets is given in Tesla (T). MRI scanners approved for human research participants have field strengths between 1.5 and 9.4T (to give you an idea of how strong such magnetic fields are: the electromagnets used to lift cars in junk yards have field strengths of approximately 1 T). Fortunately, strong magnets do not harm biological tissue which make them safe to use in human research participants (as long as participants remove all ferromagnetic objects like glasses, belts, or certain jewelry).

20.2.2 Resonance

With the vast majority of spins in the parallel state, the net magnetization in the sample points into the same direction as the magnetic field. At this point, however,

we have no chance to measure it. In order to do this, controlled changes in net magnetization need to be observed over time. This is where the resonance (the r in fMRI) comes into play: The idea behind the r is to attach energy to the nuclei which forces them leave the low-energy parallel state and flip toward the high-energy antiparallel state. This process is called excitation and is achieved by applying additional oscillating magnetic fields to the sample. It is important that the additional magnetic fields oscillate with the same frequency as the nuclei do. The spin frequency of a nucleus is called its lamor frequency. The lamor frequency depends on the amount of protons in the nucleus (which is the same in all hydrogen atoms) and the strength of the magnetic field. Because the magnet's field strength is known, the excitation signals can be adjusted to match the lamor frequency of hydrogen nuclei. As a result, energy is attached to the nuclei and they flip from their parallel spin toward the antiparallel spin. When the oscillating magnetic fields are switched off again, the nuclei will start to flip back into the parallel state while emitting the attached energy. This energy can be measured by reception coils in the MRI scanner. The emitted signal is affected by different tissue types and physiological processes. From the behavior of the nuclei returning into the parallel state, we can infer on properties of the brain tissue. Therefore, it allows for inferences in brain structure and function.

As outlined above, the nuclei precess around the longitudinal axis parallel to the magnetic field and precess in the transversal plane that is perpendicular to the magnetic field. The net magnetization (i.e., the sum of all magnetic moments) that is measured by MRI can be split up in longitudinal and in a transversal component. Without excitation of the spin system by oscillating magnetic fields, the transversal components of the net magnetization cancel each other out and only the longitudinal component parallel to the magnetic field prevails. The excitation pulses are usually designed to flip the net magnetization by 90° into the transversal plane. In consequence, the longitudinal component of the net magnetization is set to zero. As soon as the net magnetization is tipped into the transversal plane, the nuclei's precession spins will start their spins at the same starting point. In consequence, the transversal component of the net magnetization can be measured. After the excitation signals wear off, the nuclei will start to flip back into the parallel state. Two different components can be measured by the signal detection coils of the MRI scanner. First, the longitudinal component of the net magnetization will recover while the spins flip back. The longitudinal recovery is governed by a time constant that is labeled T_1. Second, the spins' coherence in the transversal plane will start to dephase until the transversal component of the net magnetization cannot be measured anymore. The transverse relaxation is governed by a time constant labeled T_2. Different tissue types (grey matter, white matter, cerebrospinal fluid, blood vessels, and bone) lead to different T_1 recovery and T_2 relaxation values. In order to construct images, spatial information must be provided along with the information on recovery or relaxation. You may recall that the lamor frequency of nuclei depends on the field strength of the magnet. Additional gradients that vary the field strength gradually across space can therefore be combined with excitation pulses at different frequencies to allow for a space dependent coding of the signal. This approach

ensures that one two-dimensional slice of the brain is measured at a time. A three-dimensional image of the brain can be mathematically reconstructed from the spatial distribution of T_1 or T_2 values across different slices.

20.2.3   From Physics to Physiology

MRI protocols that are sensitive to T_1 or T_2 contrasts provide anatomical images of the brain. To measure brain function, however, a different signal is needed. Recall that after the application of excitation pulses, the spins start to precess at the same starting point in the transversal plane, thus giving rise to the transversal component of the net magnetization. The dephasing of the spins that leads to transversal relaxation depends on interactions between the spins of nuclei. This intrinsic factor is directly reflected in the loss of T_2 signal across time (T_2 decay). Additionally, the dephasing is also influenced by an extrinsic factor. Because the spin frequency (the lamor frequency) depends on the field strength, slight inhomogeneities in the external magnetic field do also contribute to dephasing. The combination of the intrinsic and extrinsic factor leads to a signal loss in transverse magnetization that is governed by a time constant labeled T_2^*. Local inhomogeneities in the external magnetic field can depend on physiological processes in the brain. Therefore, MRI protocols sensitive to T_2^* are the backbone of functional MRI.

How do physiological processes affect the local homogeneity of the magnetic field? To answer this question we need to discuss energy consumption of the brain. The cellular basis of psychological processes can be traced to the activity of nerve cells (neurons). Neurons communicate by short transient changes of their electric resting potential across the cell membrane. This process does not rely on external energy. What does require energy, however, are housekeeping tasks of neurons such as maintaining their resting potential and restoring the resting potential after an electric signal has traveled along the cell membrane. The energy currency of the brain is a tiny molecule called adenosine triphosphate (ATP). ATP is synthesized from glucose, a sugar absorbed from food sources. This synthesis is most efficient in the presence of oxygen. Both oxygen and glucose need to be delivered to the brain via the blood stream because the brain cannot store either of the molecules. Blood is pumped through the vascular system by the heart. On its way from the heart to the brain, blood is first circulated through the lungs where oxygen is bound to hemoglobin, the oxygen transport protein in red blood cells. Then, the blood with the oxygenated hemoglobin is pumped through arteries into all parts of the body including the brain. The brain is supplied by four major arteries. After entering the cranium, arteries branch out into smaller arteries that eventually become arterioles and then capillaries. The capillaries form a fine net of tiny blood vessels that enable the exchange of oxygen, glucose, and their metabolites between the bloodstream and nerve cells. At this point, the hemoglobin trades oxygen for waste carbon dioxide and becomes deoxygenated hemoglobin.

When you put your hand onto your neck you can feel the dilatation of the arteries in response to your heartbeat. Before entering the capillaries, the pulsatile blood supply needs to be slowed down by high-resistance blood vessels to ensure a steady blood flow. Otherwise, the fine capillaries would burst from peaks in blood pressure. Where supplying arteries branch out, muscular sphincters control the blood flow into arterioles and capillaries. When nerve cells in a circumscribed region increase their activity level and thereby their energy consumption, the sphincters expand the arterioles to increase blood flow into respective regions in order to meet the temporally enhanced requirements for glucose and oxygen. That is, locally confined neural activity leads to a locally confined increase in blood flow with blood that is rich in oxygenated hemoglobin. fMRI exploits the fact that hemoglobin has different magnetic properties that depend on the binding of oxygen. Oxygenated hemoglobin is diamagnetic while deoxygenated hemoglobin is paramagnetic. Generally, objects with paramagnetic properties cause spin dephasing when introduced into a magnetic field. An increase in blood flow leads to an increase in oxygenated hemoglobin relative to deoxygenated hemoglobin. In turn, it leads to less spin dephasing and in consequence to slower transversal relaxation and a stronger T_2^* signal. That is, MR protocols sensitive to T_2^* can use oxygen as an intrinsic contrast agent of the brain for the mapping of neural activity. In this case, we speak of blood oxygen level dependent fMRI, or in brief, of BOLD fMRI.

Because the recorded signal in fMRI relies on blood flow dynamics in response to changing neural events, the signal is called the hemodynamic response. The typical hemodynamic response as revealed by BOLD fMRI starts with a temporal offset of 1–2 s to the neural activity that triggered the response. The reason for this time lag reflects the time window until the feedback loop between active neurons and their supplying blood vessels has increased the local blood flow. After a steep rise, the hemodynamic response peaks about 4–5 s later and then falls steadily over another 5–6 s until it falls below baseline 12–13 s after the triggering neural activity,. The BOLD signal returns to baseline level approximately 20 s after the onset of the neural events. From this timing information, we can see that the hemodynamic response lags the neural events behind and is rather slow compared to psychological processes that often take only a couple of hundred milliseconds to finish. Nevertheless, neural correlates of even short-lived psychological processes can be traced by BOLD fMRI, given individual experimental trials are sufficiently spaced.

The sampling rate of the MRI scanner needs to be set in a way that sufficient information on the hemodynamic response will be acquired. The time between successive excitation pulses of the scanner is called repetition time (TR) and quantifies the acquisition speed of the scanner in a given experiment. Modern fMRI scanner that uses echo planar imaging (EPI) pulse sequences can scan the majority of the brain with a TR of 1–2 s while retaining a sufficient spatial resolution (usually 3 mm^3 voxels). It has been demonstrated that sampling rates below 0.5–1 Hz do not substantially improve the measurement. That is, a TR of 1–2 s provides an appropriate temporal resolution for fMRI even if the examined psychological processes follow a faster time scale.

Fig. 20.2 A high-resolution T1-weighed anatomical MRI scan and a BOLD fMRI image (T_2^*-weighed) from the same participant
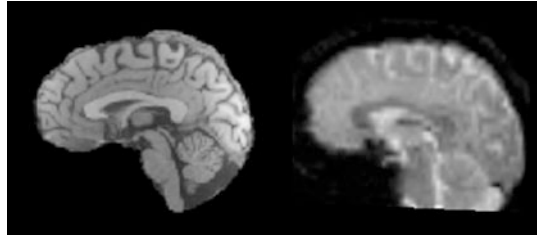
The EPI pulse sequences acquire data in two-dimensional slices from which a three-dimensional image of the brain can be reconstructed. The brain spans about 12.5 cm from the brainstem to the most dorsal part of the parietal lobe. With 3-mm-slices spaced by a 0.3 mm gap, it would require 38 slices to image the entire brain. Because the TR depends critically on the number and the spacing of these slices, the temporal resolution can be improved by omitting parts of the brain during image acquisition, in most cases the brain stem and the cerebellum. This is a feasible approach especially in neuroeconomical studies, because cortical or midbrain structures lie in main interest of most investigations. In the following, we will refer to functional images acquired in one TR as "volumes" and not "brains" to emphasize that the processing steps are applied to the functional data irrespective of the degree to which the entire brain is covered during imaging. Figure 20.2 shows a T_1-weighted anatomical and a T_2^*-weighted functional volume from the same participant.

20.2.4 Summary

In the first part of this chapter, we have established that fMRI measures magnetic properties of brain. Furthermore, we have discussed how vascular activity in the brain gives rise to the blood oxygen-dependent signal that can be measured by MRI scanners and allows for inferences on neural activity. We have concluded with remarks on temporal and spatial properties of the hemodynamic response. This part was supposed to give a brief overview on the physical and physiological basis of fMRI. For more in-depth information on the physical and physiological basis of fMRI, we refer to the excellent textbook by Huettel et al. (2009). In the following, we will deal with the statistical analysis of the acquired volumes in the context of statistical parametric mapping.

20.3 Analysis of fMRI Data

The analysis of fMRI data can be separated into three consecutive steps: (a) pre-processing of functional images (b) first-level analysis of fMRI time series, and (c) second-level (or higher order) analysis. Preprocessing describes necessary

analysis steps that are carried out to ensure that all data are in the same five-dimensional coordinate space which is a prerequisite for statistical analysis (we will explain the five dimensions in the following). Then, the first-level analysis is carried out separately for each participant. It outputs statistical parameters that are eventually fed into the second-level analysis that aggregates the data across participants for statistical inference on activation patterns between groups or in the population the sample of study participants has been drawn from. All three analysis steps can be carried out in freely available analysis software tools. The two most popular tools in the neuroimaging community are Statistical Parametric Mapping (SPM) that is issued by the Wellcome Trust Centre for Neuroimaging (http://www.fil.ion.ucl.ac.uk/spm/) and FSL, issued by the Oxford Center for Functional Imaging of the Brain (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/). Both imaging software packages can be downloaded for free from the respective websites and come with detailed documentation and example data sets.

Before we start to discuss preprocessing and first- and second-level analyses in more depth, we start with some general remarks on experimental design, as this is the prerequisite to understand what is going on during the analysis steps.

20.3.1   Experimental Design

One inherent property of the BOLD signal is that it is no absolute signal: We always need to compare the signal to some sort of baseline or control condition. This control condition can be implicit, that is the hemodynamic activity evoked by a task is compared to hemodynamic activity while there is no task. This, however, might come with the downside that the task and the control condition differ in many different aspects. Let us assume we are interested in hemodynamic activity evoked by the feedback about the second mover's behavior in the trust game. In the experimental condition, the research participants in the role of the proposer face an information screen that states whether the responder is defected or not. To ensure that the participants actually process the information on the display, they are asked to execute some sort of manual response to the information. The evoked hemodynamic response could be compared to a condition were the participants did nothing. Such a condition, however, would not only differ in the decisive variable (cooperation versus defection) but also in physical appearance of the display, the lack of a motor response, and the absence of a monetary outcome. Thus, hemodynamic activity associated with these factors cannot be easily disentangled from the actual activity inherent to the trust game. Therefore, it is a better idea to contrast the task condition with an explicit control task that differs only in one aspect critical to the study. In our example, this could be a computer-raffled lottery where the participants either loose or win but that most critically lacks the social component of the trust game (for example, see Delgado et al. 2005). This experimental design relies on the pure insertion principle inherent to subtractive experimental methods: Different processes are assumed to be additive. By subtracting the hemodynamic

response during one process (computer lottery) from the hemodynamic response during another process (outcome of the trust game), the mere difference in hemo-dynamic activity between both processes survives the subtraction (neural response to cooperation or defection). This approach can be very powerful and useful, it disregards, however, the possibility of non-additivity in the sense of interactions between task conditions. One way to study such interactions is the use of factorial designs in which all combinations of two or more independent variables with two or more levels are administered. If, for example, one would want to study the effect of absolute versus relative income, possible independent variables are absolute income (receiving a high or a low amount of money) and relative income (receiving more, less or just as much as somebody else). In consequence, six different conditions arise that allow to disentangle additive and interactive effects of absolute and rel-ative income on reward related hemodynamic activity (see Fliessbach et al. 2007, for example).

A further comparison strategy is parametric designs. In parametric designs, the covariation of the BOLD response with a parametrically manipulated independent variable is examined. If, for example, we are interested in neural correlates of decision utility during gambles, we could vary the possible gains associated with the gambles parametrically and examine whether the BOLD signal in a given brain region responds contingently (see for example Tom et al. 2007).

The previous considerations all dealt with comparison strategies. A further thriving issue concerns the temporal sequence of stimulus presentation. There are two main approaches to stimulus timing in BOLD fMRI experimental design: blocked versus event-related presentation. In a blocked design the research par-ticipants are asked to alternate between blocks of many trials in the experimental and the control task. Ideally, the length of the blocks corresponds to the length of the hemodynamic response (about 10 s). With shorter time intervals, the BOLD response cannot return to its baseline and the differences between experimental and control conditions become blurry. With longer blocks, on the other hand, scanner drift can inflate the differences between experimental and control blocks by intro-ducing noise to the data. In blocked designs, many experimental trials contribute linearly to the recorded hemodynamic response. Therefore, blocked designs come with a high power to detect differences in activation between conditions. The downside, however, is that temporal information on the hemodynamic response cannot be analyzed. Furthermore, many research questions cannot be operational-ized in blocked designs. Many research designs do not allow for an a priori specification of task and control conditions. Let as assume we are interested in examining hemodynamic activity associated with continued gambling to recover previous losses, a phenomenon called "chasing losses", which is maladaptive decision behavior common in pathological gambling (Cambell-Meiklejohn et al. 2008). Participants are confronted with the outcomes of gambles and decide whether they want to continue or quit gambling after experiencing losses. In such an experiment, the participants' decision behavior determines if a given experi-mental trial is assigned to the task (chasing losses) or to the control condition (quitting gambling). If this research question would be addressed in a blocked

design, the researchers would need to tell participants to chase losses in one block of trials and to quit gambling in others. This, however, would eliminate the critical behavior under the researchers' scrutiny: Participants would not make the decision to continue gambling by themselves and no neural activity associated with the decision making process could be recorded. Because decision-making is one of the main research interests in neuroeconomics, most studies in the field adopt event-related designs. In event-related designs, the evoked hemodynamic response to single experimental trials is examined. The advantage of event-related designs is that events can be assigned to experimental conditions post hoc. It is also possible to exclude certain trials from the analysis, for instance error trials or trials in that the participant fails to respond in a given time window. Furthermore, event-related designs allow for a precise temporal characterization of the hemodynamic response. Compared to blocked designs, however, they lack a high degree of detection power. A further potential drawback is that subsequent presentation of the same events can introduce an artificial blocked design, where the BOLD response saturates and becomes equivocal to different task conditions. To counteract this problem, the interstimulus interval (ISI) should be large and jittered which means that it is randomly varied in its duration across experimental trials (e.g. a randomly chosen ISI between 3500 and 6000 ms).

Now that we have established, how experiments can be designed to be suitable for fMRI we can go on with a discussion how to analyze the data.

20.3.2   Preprocessing

As we have discussed earlier, the best way to think about fMRI data is in voxels. Each functional volume lies within a three-dimensional grid comprising a large number of voxels with one activation value assigned to each of them. Each fMRI run comprises many of these volumes which are acquired consecutively over time. This results in an activation time series for each voxel (see Fig. 20.3) and, if we look at all voxels and all time-points at once, one-four dimensional data set (x by y

Fig. 20.3 The BOLD time series as measured from one gray matter voxel over a time period of approximately ten minutes
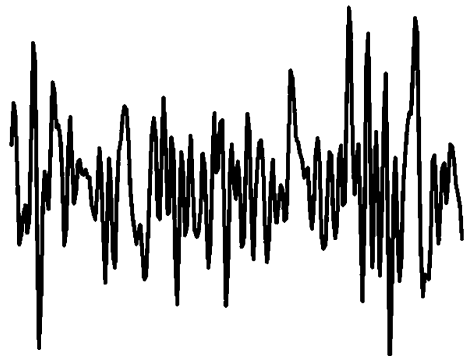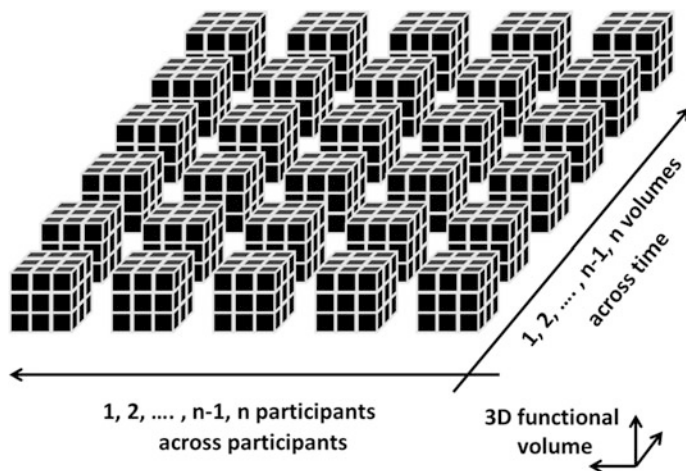
Fig. 20.4 The five-dimensional nature of BOLD fMRI experiments. The first three dimensions refer to the three-dimensional brain that is scanned across time (dimension 4). The fifth dimension refers to different participants who undergo the same imaging protocol

by z by time). We acquire such a four-dimensional data set for each participant in our experiment. That is, we can think of the data of an entire experiment as a five-dimensional matrix (a three dimensional brain scanned across time and across participants, see Fig. 20.4).

Now imagine a perfect world: We can expect various things from a perfect world: First, we would expect that all voxels in one volume were sampled at the same time (that is, at a given point in time, for a given participant, all data along the first three dimensions of our matrix are acquired simultaneously). Second, along the time dimension, we would expect that data measured in the same voxel merely reflects changes in activation (and not sampling error) at the same precise anatomical location (and not that of a neighboring location). And at last, we would expect the brains of all participants to match spatially, that is, we would expect that along the fifth dimension (the participants) a given voxel always corresponded to the same neural structure. Alas, the world is not perfect, especially not for neuroimagers: The EPI pulse sequences used for fMRI measure a volume one slice at a time. With a TR interval of two seconds, this implies that the temporal offset of two voxels in a single volume can be as large as two seconds. Furthermore, even the best research participants with the highest motivation to keep still during the experiment will move their heads no matter how tightly we constrain them mechanically. Along the time dimension, every millimeter of motion will move consecutive volumes further away from the first volume, slowly but dramatically distorting our data across time. Finally, we can intuitively agree that our expectation on the uniformity of brains across participants cannot hold: It is not only that peoples' heads and brains vary in size. There are also individual differences in the brains' gyri and sulci.

Given these constraints on our data, we need to come up with an idea to resolve these issue. Otherwise we would not be able to meaningfully analyze our data. Fortunately, there are powerful algorithms available to correct for temporal and spatial distortions within single volumes, across time and across participants. However, we should be aware that we substantially alter our data during preprocessing and that in consequence, statistical inferences on brain activation are not carried out based on data as they were initially measured. Therefore, all preprocessing should be applied carefully and with high caution. Preprocessing steps are nevertheless necessary means to correct for deviations from our perfect neuroimaging world: They ensure that subsequent statistical analysis will be meaningful.

Usually, preprocessing steps include (1) slice timing, (2) head motion correction, (3) coregistration/normalization, and (4) spatial smoothing. Temporal bandpass filtering and detrending can be applied as additional but not necessary steps.

Slice timing refers to the correction of acquisition delays between slices within the same volume. For the slice timing correction, the user specifies a reference slice for each volume. Then, an algorithm analyzes the time course in each voxel across volumes and interpolates how the data points in every other slice of a given volume would have looked like if they would have been acquired at the same time as the reference slice. The further away a given slice lies from the specified reference slice, the stronger the data are altered by the algorithm. Therefore, it is a wise idea to choose the slice in the middle of the volume as reference. In many cases, the slices of a single volume are not acquired in ascending or descending order but interleaved (that is, all even slice numbers are imaged before the odd ones). In this case, slice timing correction is always recommended. In cases where the scanner has acquired the slices consecutively, slice timing can be omitted if steps are taken to control for slice acquisition offsets later on during statistical analysis (and the TR is not too large).

Motion correction provides an algorithm to correct for spatial distortions across volumes within a single participant because of head motion. Fortunately, we can be sure that our participants' heads do not change their size or form during the relatively short fMRI run. That is, all gross deviations between heads across volumes are almost exclusively attributable to head motion. In the scanner, a head can move up and down, from left to right and back and forth. That is, it can translate along three dimensions. Furthermore, it and can rotate around three axes (nodding, shaking and tilting). We can therefore align the heads within all subsequent volumes with the head in the first volume with a six-parameter transformation. Because the head itself does not change during transformation, we call this a rigid-body transformation which has six degrees of freedom. Usually, the parameters for each volume are saved so that they can be used as covariates (nuisance regressors) during later statistical analysis. The motion parameters should also be examined for outliers. If single participants have moved too excessively during scanning it might be wise to exclude them from further analysis (usually, translations less than three millimeters and rotations less than three degrees are considered tolerable).

Coregistration and normalization are two consecutively applied preprocessing steps that ensure comparability of the data across participants. As mentioned earlier, peoples' brains to not only differ in size but also in gyrification and sulcification. Therefore, a rigid body transformation as discussed above or a linear transformation are both not suitable to match one brain with the others. The solution is an affine transformation with twelve degrees of freedom in total. That is, twelve parameters are needed to match an individual brain with a group. This process is called normalization. One way to do this would be to choose one representative participant from your sample (maybe the one who is closest to the mean of demographics like age or education and a member of the more frequent sex). The next step would be to apply the affine transformation. During this process all other brains are resized, squeezed and dragged until they match the reference brain most closely. After the transformation, each voxel in each volume will contain information on the activation level of exactly the same neural structure across participants. This approach of choosing a representative reference brain from the study sample, however, would come with two major downsides. On the one hand, the researcher would be in need of a high degree of anatomical skill to precisely identify what structures are activated during later analysis. On the other hand it would be quite tricky to compare results across different experiments and across publications. For that reason, neuroimagers have agreed on a standard reference brain in a standard coordinate space. This standard brain has been issued by the Montréal Neurological Institute (MNI). It unifies anatomical information from 152 white individuals and should therefore be a close reference template to most brains from healthy individuals of European descent. Other reference brains are available for individuals from other populations. Practically, we need to deal with one problem during normalization: As you can see from Fig. 20.2, the spatial resolution of functional EPI images is not as good as the spatial resolution of anatomical T_1-weighted scans. If the normalization algorithm was fed by anatomical information from the functional volumes alone, it would lack detailed information. Coregistration prior to normalization is a recommended resolution for this problem. Coregistration takes advantage of two facts: First, it is relatively easy to precisely align a high resolution anatomical T_1-weighted image with functional volumes from the same person because it requires only a rigid-body transformation (see above). Second, it is also relatively easy to normalize a T_1-weighted image of a study participant to a T_1-weighted reference brain in standard coordinate space because of the high anatomical detail of the individual T_1-image. For this purpose, a high resolution T_1-weighted anatomical image is usually acquired along with the functional volumes. Then, the functional volumes are aligned with the individual anatomical image. In a second step, the individual anatomical image is normalized to the reference brain and the resulting twelve transformation parameters are then applied to every single functional volume. In easier words, the individual anatomical image "piggybacks" the individual functional volumes to obtain the best normalization results possible.

Spatial smoothing is the last strongly recommended preprocessing step. The rationale behind smoothing is to use the information of neighboring voxels to smooth the signal from each voxel in a volume. As a result, the images become

more blurry, but also increase in their signal-to-noise ratio. Smoothing leads to an increase in statistical power of subsequent statistical tests because the error terms of the test statistics are reduced and because single peak activation foci will more likely merge to robust activation clusters. Smoothing takes advantage of the fact that the time courses of adjacent voxels are highly intercorrelated because for most psychological processes, the spatial resolution of fMRI scanners exceeds the functional resolution required to image their neural correlates. Testing theory states that each measurement is additively composed of a true value and an error term. The error terms are thought to be independent from each other and have an expected value equal to zero. Thus, if we average the signal measured in neighboring voxels, we effectively decrease the error term while the true values are more or less left as they are. The averaging procedure during smoothing applies weighting to neighboring time series in a way that the time courses of more nearby voxels contribute more strongly. The weighting is accomplished by applying a three-dimensional Gaussian kernel that has its peak on the voxel to be smoothed. The size of the Gaussian kernel is given by its full width at half maximum (FWHM). The recommended size depends on the resolution of the functional data (and the neural structures and psychological processes that are imaged). In most cases, the kernel size varies between 6 and 12 mm FWHM.

This concludes the strongly recommended steps. Additional preprocessing can include detrending and temporal bandpass filtering. Detrending corrects for linear trends apparent across the entire time series because such linear trends are most likely attributable to MRI scanner drift. Bandpass filtering removes frequencies in a specified frequency band from the time series because certain frequencies do not reflect neural activity and are therefore most likely attributable to scanner artifacts or introduced by cardiorespiratory activity of the research participants. In most cases, a high pass filter of 0.008 Hz is applied, that is, all oscillations in the frequency bands below this threshold are removed from the data.

20.3.3   First-Level Analysis

A typical first-level analysis aims at the isolation of activation differences between our experimental conditions of interest on the level of single participants. In consequence, the first-level analysis is carried out for each participant separately. The most common approach in the context of statistical parametric mapping is to set up a statistical model that explains the acquired neurophysiological data best and then conduct inferences on activation differences based on the parameters from this model. In simple words, the activation data are correlated with the temporal sequence of experimental events voxel by voxel and inferences are subsequently conducted based on the correlation coefficients. Usually, a mass univariate approach is applied: Parameter estimation and statistical inference are conducted separately for each voxel and results are only combined in the very last step.

20.3.3.1   Model Specification and Parameter Estimation

The statistical model most widely applied in neuroimaging is called the general
linear model (GLM). In a GLM analysis, a dependent variable (in our case the
neurophysiological data) is predicted by a set of predictors that are linearly (additively) combined. In the most simple case, the model equation of a GLM is
$y = mx + n$, where y refers to the dependent variable, x to one predictor, m to a
weight attached to the predictor (the model parameter) and n to an intercept (a
constant value added to the equation). You may notice that this equation resembles
the linear equations you have solved in high school. Further predictors can be added
to the model (like $y = mx_1 + px_2 + zx_n + n$) if we believe that this leads to a better
prediction of y. In the context of fMRI, the dependent variable is the BOLD time
series from one voxel. This time series has as many entries as there are volumes in
our fMRI run (in the following we will refer to these time series as vectors). The
model equation needs to be designed in a way that the predictors and their corresponding weights output values for y that come closest to the values in the time
series vector. The experimental conditions in the experiment serve as predictors
with one predictor for each condition. In the simplest case one experimental condition (for instance tapping with the right index finger) is compared with a control
condition (doing nothing; most critically, no tapping with the right index finger).
A basic GLM for this design would be $y = mx_{\text{tapping}} + px_{\text{nothing}} + n$. Let us assume
that the participant in this example experiment alternated between the two tasks
(tapping versus doing nothing) every 20 s for seven times in total. This experiment
would last for 280 s. If we choose to image the brain at a TR of two seconds, we
would acquire 140 functional volumes in the experiment, leaving us with a vector
with 140 entries per voxel. These 140 observations per voxel serve as the criterion
variable y. What do the predictors look like? As mentioned earlier, we plan to
correlate the neurophysiological data with the temporal sequence of experimental
events. Therefore, we need vectors for each predictor that have as many entries as
the BOLD vectors holding the criterion variable. In the simple case of our example,
we will code the onsets of the experimental condition with 1 and leave the value to
1 for the entire duration of the experimental block. That is, the vector of the first
predictor will hold a 1 whenever the participants performed on the tapping task and
a 0 whenever they did nothing. Thus, the first 10 entries of the vector will hold
ones; the next 10 entries will hold zeros and so on. The onset vector for the second
condition (resting) will hold ones whenever the participants are resting, and zeros
whenever the participants perform on other tasks (in our case, tapping their right
index fingers). Now that we have modeled all experimental conditions, further
predictors could be added that are of no interest for the research question per se but
might enhance the model fit by explaining systematic noise in the data. For
example, we could use the six parameters from the motion correction during preprocessing as six additional predictors. Neuroimagers refer to these control predictors as covariates of no interest or as nuisance regressors. Note that we are still
dealing with the data of one participant. Therefore, nuisance covariates such as age
or gender that vary across participants cannot be entered to the model at this point.

After the specification of all predictors, we end up with an $m \times n$ matrix, where m refers to the number of functional volumes and n to the number of predictors. In this simple case, we have coded our experimental conditions in a binary mode (that is we have only used ones and zeros). If we believe that the BOLD signal increases stronger in different realizations of our predictor variable (for example, we measure brain reactivity to economic gambles and believe that BOLD activity increases with increasing gains), we could modulate the predictor's vector with a parameter. In the example with the gambles, one predictor would carry ones and zeros (whenever the gamble was shown) and a further predictor would carry the modulator (the potential gain of the gamble as an integer) whenever the former vector carried a 1.

There is reason to believe that the BOLD signal does not rise and fall in a way the binary predictors suggest. Figure 20.5 depicts a function that describes the hemodynamic response that is usually measured in response to stimulation. To achieve a better fit between the empirical data and the predictors from the GLM, the onset vectors are convolved with such a hemodynamic response function (HRF). In most cases, a canonical HRF is used that is distributed along with the analysis software packages. A further widely applied option is to additionally convolve the predictors with the HRF's first temporal derivative. This additional step is a good means to account for temporal variability of the onset of the hemodynamic response. In case that slice timing is omitted during preprocessing, this method is highly recommended.

After the GLM is specified and all regressors in the equation (experimental conditions, parametric modulators and nuisance covariates) are convolved with the canonical HRF (and its temporal derivative) the regression weights (called beta
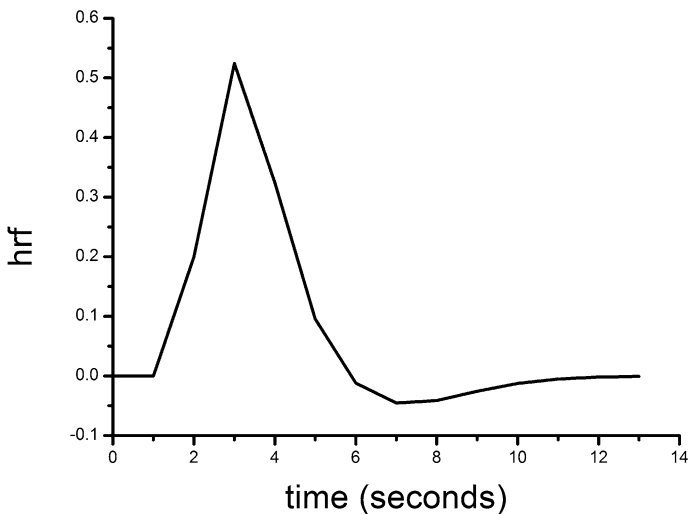

Fig. 20.5 The canonical hemodynamic response function as distributed alongside the SPM software package

weights) need to be estimated. This is done by a least square approach: The beta weights are set in a way that the sum of the squares of the deviations between the predicted values (the y values) and the empirical data (the BOLD time series) is minimized. At the end of this process, we end up with one beta weight for each predictor for each voxel. Call yourself in mind that we are still dealing with single subjects. The process of setting up the GLM and estimating the corresponding beta weights is repeated for each participant in the sample. In our simple example experiment, the temporal layout of experimental conditions is essentially the same for all participants; i.e., the same GLM with exactly the same predictors (except for nuisance covariates that are specific for individuals) can be used. In case of more complex experimental designs, (for instance an event-related design with events presented in random order), the GLM itself (that is the columns of the $m \times n$ matrix) would be the same, the onsets (and modulators), however, would be different for each participant.

20.3.3.2 Contrasts

Now that we have ended up with beta weights, we can conduct statistical inference on these weights. This is still carried out separately for each participant. The question addressed by first-level inferences is: In which voxels and to what extend do given experimental conditions lead to activation differences? In the most simple and most widely used case this question is answered by a t-test. A t-test is a statistical test that contrasts two measurements and assesses whether the difference is significantly larger as the general variability in the data set. In the case of a first-level inference, the t-test looks if the difference between the beta weights of two experimental conditions is larger than an error term that is calculated from the variability of the fMRI data the GLM cannot explain (while also taking the number of observations and predictors in the model into account). In more formal terms, the t-test outputs a t-value that is a quotient of the difference between the beta weights divided by this error term (sum of squared residuals of the GLM minus number of functional volumes minus number of all predictors in the GLM). These t-values follow a t-distribution. Because t-distributions are well-known probability distributions, we can look up the probability by which the obtained t-statistic suggests a difference between the beta values even though there is not one. If this probability is sufficiently low, we conclude that the beta weights differ significantly. Different t-distributions differ from each other in their degrees of freedom (df). For first-level inferences, the df that correspond to the t-test can be calculated from the number of data points (number of functional volumes) minus the number of predictors in the model.

Let us consider our example experiment once again: We would definitely want to look where activity increases during finger tapping as compared to the resting condition. For the first contrast, we would therefore subtract the beta value of the resting-regressor from the beta value of the tapping-regressor for each voxel and assign this difference to the nominator of the t-value. If we would also be interested

in activity decreases during motor tapping, we would calculate a second contrast for which we subtract the beta value of the tapping-regressor from that of the resting-regressor. The error term is essentially the same for both contrasts, but differs slightly across voxels: For each data point in our dependent variable we would calculate the difference between the actual BOLD activity and the value predicted by our GLM (the residuals). These differences are squared to control for different signs and then summed up across data points. Finally, we would subtract the number of data points (240) and the number of regressors in the GLM (two conditions plus six motion regressors of no interest equals eight) from the sum of squared residuals. This whole term would then be assigned to the *t*-value's denominator. This step is repeated for all voxels in the brain and then for all participants in our sample. Thus, we would end up with one map per contrast and participant that maps the *t*-values on the brain. Finally, we would calculate the amount of degrees of freedom for our statistical test (240 data points minus eight regressors equals 232 degrees of freedom) and look up the corresponding probability values (*p*-values) to each *t*-value that tells us the probability by which the difference in the beta values can be attributed to chance. Because we want to be sure to not assume a significant difference between experimental conditions when there is none, we would want this probability to be low, for example below 0.001 %. With a *t*-distribution with 232 df, the *t*-value would need to be as high as 1.651 to be considered significant.

This, however, would us only leave with an assessment of statistical significance between experimental conditions within subjects. In order to assess significance across or between subjects, we would need to perform additional analyses.

20.3.4   Second-Level Analysis

The rationale behind the second-level analysis is to combine the results from the first-level analyses of all participants and assess statistical significance across participants. Various statistical inferences are possible: In the simplest case, one would want to examine whether neural activity differences between the conditions observed on the first-level are idiosyncratic to single participants or can be found in the majority of participants. This could be accomplished by a one-sample t-test. A one-sample t-test tests whether the mean of a dependent measure that is calculated across participants differs significantly from a given value. In the context of a second-level fMRI analysis, this t-test looks if the mean of activation in one voxel is significantly different from zero. Different statistical models can be set up, depending on the research question asked: If, for instance, a group of pathological gamblers is compared with a group of healthy control participants regarding their neural response to risky choice, a two-sample t-test that tests for differences in the mean of neural activity between the two groups is the appropriate statistical test. Multiple linear regressions and multifactorial models are also possible for more complex experimental paradigms. Please note that in the vast majority of cases,

second-level analyses model participants as random effects. This statistical approach assumes that participants were sampled from an underlying population and ensures that any inference drawn from the data can be generalized to this population.

Up to this point, we have not specified on which values the second-level analysis is carried out. The first-level analysis outputs a t-value for each voxel. This t-value, however, is not the best means to quantify activation differences between experimental conditions. As outlined previously, the t-values depend not only on the difference in the beta values but also on the number of data points (i.e., functional volumes) and predictors in the model. It is a better idea to use a standardized measure to quantify activation differences. Such a standardized measure is the effect size that can be calculated for each statistical test on the first level. The effect size quantifies the magnitude of the activation differences in units of measurement and is fed into the statistical tests on the second level. For that purpose, the mean (or differences in means) and the variability of the effect sizes are calculated and compared in a term such as the t-term. The corresponding df for the t-distribution can be calculated from the sample size and the number of cells in the experimental design.

For the second-level analysis, we are more interested in assessing whether an observed activation difference in our sample reflects a true relationship in the population than in estimating the size of the effect. Therefore, the statistical parameters mapped for the visualization of results are usually the test statistics (such as t-values) rather than the effect sizes of the second-level tests. The test statistics are thresholded at a given probability (e.g., $p < 0.001$) which means that only parameters are considered that are high enough to let us assume that the observed activation difference cannot be attributed to chance. For visualization purposes, the statistical parameters that survived the statistical thresholding are color coded and projected onto a structural MR image in standard coordinate space. Besides from this height threshold, an extend threshold can be applied additionally. Because it is very likely that the statistical parameter of a single spatially isolated voxel exceeded the threshold by chance, it is recommended to specify a minimum cluster size in voxels (e.g., $k > 8$) and ignore clusters of adjacent voxels if the number of co-jointly activated voxels does not exceed this minimum cluster size.

At the end of all these processing steps (preprocessing, first- and second-level analysis including thresholding) stands one statistical parametric map that informs us about activation differences between experimental conditions. However, this approach comes with one major downside that will be discussed and resolved in the next paragraph.

20.3.4.1 Multiple Comparison Correction

In 2009, a wave of gloating newspaper articles was published in the popular media sarcastically criticizing functional MRI. The opinions were based on an a study that had been presented at the Annual Meeting of The Organization for Human Brain

Mapping (OHBM) earlier that year. What had happened? In the study that is available online as a conference poster (Bennett et al. 2009) the authors had put a dead Atlantic salmon (originally bought for food consumption) into the scanner and run structural and functional MRI on it. While the salmon was in the scanner, the authors presented visual stimuli and "asked" the salmon to perform a task. Even though the salmon did not respond to the task behaviorally (most certainly because it was dead), the authors still ended up with BOLD time series and temporal information on the sequence of task events that was eventually correlated with the neurophysiological data. The authors set up a GLM and processed the data up to the first level (because there was only one salmon involved) as discussed earlier, calculated a contrast between experimental conditions and thresholded the resulting t-map at $p < 0.001$ with an extend threshold of three adjacent voxels. Given that the salmon was not alive during scanning and the task more suitable for human research participants, results were strikingly surprising: Significant clusters lit up in both the salmon's brain and its spinal cord! Clearly, these results were easy bait for the popular media and fueled skepticism toward neuroimaging in general. How can we trust neuroimaging results in human research participant when even a dead fish shows brain activity which looks like it was evoked by an experimental task? To answer this question we need to take a closer look at a problem called the multiple comparison problem and how it can be resolved in the context of fMRI.

As outlined in the previous paragraph, the rationale behind statistical hypothesis testing is to calculate a probability by which the empirically observed difference (or relationship) in the data was obtained given that there is no such relationship in reality. If this probability is sufficiently low (e.g., below 0.001 %), we conclude that the observed effect must reflect a true relationship. However, we should be aware that in one out of 1000 cases, this conclusion would be wrong—simply because a very low probability only makes things become unlikely but does not rule them out entirely. A typical fMRI volume consists of ten-thousands of voxels. Because the analysis is run separately for each voxel, the number of statistical tests is as high as the number of voxels. Across all tests, the probability of erroneously assuming a true relationship because of a very low p-value increases dramatically. Such errors are called false positives (or alpha-errors) and describe a situation where we decide to accept a statistic as evidence for a difference between experimental conditions even though there is none in reality. The problem that arises from the massive amount of statistical tests run during fMRI analyses is called the multiple testing problem, or alpha error inflation. This is what happened in the dead salmon study: Because of the mere number of comparisons, some voxels lit up by chance because the high t-values suggested an activation difference even though the salmon was not paying attention to the task (which we can infer from the behavioral data and the fact that it was dead). Luckily for the neuroimaging community, there are various methods that can be used to eliminate (or at least minimize) the multiple comparison problem.

When we conduct statistical inferences on second-level data we want to make sure that we do not erroneously assume that a single voxel is active even though it is not. Furthermore, we want to make sure that across all statistical tests, the chance of

obtaining a false positive result is very low as well. We can ensure this by correcting for the number of tests conducted. Because we control the alpha error probability for a family of statistical tests, this procedure has been labeled family wise error (FWE) rate correction. The standard method for this is the Bonferroni correction, which accepts a family of statistical tests (such as in the entire brain) as significant if the alpha error probability of each test in the family is below a specified significance level that is divided by the number of tests in the family. That is, if we consider a false positive rate of 0.05 % as acceptable and there are 10,000 voxels in our data set, we would consider a voxel active if the error probability of its t-value fell short of $0.05/10,000 = 5.0e - 6$. Applying the Bonferroni procedure effectively minimizes the risk of false positive results but it comes with a major downside. It is very conservative: Whenever we conduct a statistical test, there is not only the risk of committing an alpha error (i.e., assuming a difference even though there is none in reality) but also the risk of committing a beta error (i.e., assuming no relationship although there is one in reality). Both types of error depend on each other: With an increasing statistical threshold, the probability of committing an alpha error decreases but beta errors on the other hand become more likely. Therefore, Bonferroni correction admittedly controls for alpha error commission but it also increases the chance of committing a beta error. In the context of fMRI there is reasonable doubt that the Bonferroni correction is the gold standard for multiple comparison correction. The Bonferroni correction is appropriate for families of independent statistical tests. In fMRI data, however, full independence between all tests cannot be assumed: As outlined earlier, data points from neighboring voxels are highly intercorrelated and these correlations are further amplified by spatial smoothing during preprocessing. Therefore, the number of test families that require correction is substantially lower than the amount of voxels in the brain, leaving Bonferroni correction as a too conservative approach. The more appropriate routine is the application of Gaussian random field theory (RFT) to control the family wise error rate. RFT is a research body in mathematics that deals with smooth statistical maps (such as our t-maps). RFT provides tools to estimate the overall smoothness of the fMRI data set which depends on the degree of spatial correlations between voxels in the raw data and the size of the Gaussian kernel applied during smoothing. If the overall smoothness is known, the number of resolution elements (resels) in the data set can be calculated. The number of resels equals the number of independent observation in the data set and gives the amount of test families for that we need to correct. From the number of resolution element, we can determine the expected Euler characteristic (EC). In the context of functional imaging, the EC gives the number of expected clusters in a smooth statistical map after thresholding. Because the expected EC depends on the statistical threshold, it is approximately equivalent to the probability of committing a family wise error. Thus, the statistical threshold for the second-level analysis that corrects for multiple comparisons by controlling the family wise error can be inferred from the EC. This method has been implemented in most analysis software packages. To come back to the dead salmon in the scanner: After controlling the FWE according to Gaussian RFT no active voxel could be observed in its central nervous system.

When we ask how we could trust neuroimaging results when even dead fish show neural activation contingent with experimental tasks, the answer clearly is: When we control for multiple comparisons!

20.4 Conclusion

The purpose of this chapter was to give an overview on how fMRI works. We have covered the physical and physiological basics of the BOLD signal and discussed processing steps and statistical analysis of BOLD time series in the context of statistical parametric mapping. It would have been beyond the scope of this chapter to discuss more advanced functional imaging methods than the mass univariate approach which is the most common way to analyze fMRI data. For more in-depth information on statistical parametric mapping, multivariate approaches to BOLD fMRI data and the analysis of functional connectivity between different brain regions during task performance, we would like to refer to the SPM textbook by Friston et al. (2007) and the textbook by Huettel et al. (2009).

References

Bennett CM, Miller MB, Wolford GL (2009) Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: an argument for multiple comparisons correction. Neuroimage, 47(Suppl 1):S125.

Campbell-Meiklejohn DK, Woolrich MW, Passingham RE, Rogers RD (2008) Knowing when to stop: the brain mechanisms of chasing losses. Biol Psychiatry 63(3):293–300

Delgado MR, Frank RH, Phelps EA (2005) Perceptions of moral character modulate the neural systems of reward during the trust game. Nat Neurosci 8(11):1611–1618

Fliessbach K, Weber B, Trautner P, Dohmen T, Sunde U, Elger CE, Falk A (2007) Social comparison affects reward-related brain activity in the human ventral striatum. Science 318 (5854):1305–1308

Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ (1995) Statistical parametric maps in functional imaging: a general linear approach. Hum Brain Mapp 2(4): 189–210

Friston KJ, Ashburner JT, Kiebel SJ, Nichols TE, Penny WD (2007) Statistical parametric mapping: the analysis of functional brain images. Elsevier, Oxford

Huettel SA, Song AW, McCarthy G (2009) Functional magnetic resonance imaging. Sinauer Associates, Incorporated

Tom SM, Fox CR, Trepel C, Poldrack RA (2007) The neural basis of loss aversion in decision-making under risk. Science 315(5811):515–518