Visualization of Group Inference Data in Functional Neuroimaging

Jan Gläscher

Abstract While thresholded statistical parametric maps can convey an accurate account for the location and spatial extent of an effect in functional neuroimaging studies, their use is somewhat limited for characterizing more complex experimental effects, such as interactions in a factorial design. The resulting necessity for plotting the underlying data has long been recognized. Statistical Parametric Mapping (SPM) is a widely used software package for analyzing functional neuroimaging data that offers a variety of options for visualizing data from first level analyses. However, nowadays, the thrust of the statistical inference lies at the second level thus allowing for population inference. Unfortunately, the options for visualizing data from second level analyses are quite sparse. rfxplot is a new toolbox designed to alleviate this problem by providing a comprehensive array of options for plotting data from within second level analyses in SPM. These include graphs of average effect sizes (across subjects), averaged fitted responses and event-related blood oxygen level-dependent (BOLD) time courses. All data are retrieved from the underlying first level analyses and voxel selection can be tailored to the maximum effect in each subject within a defined search volume. All plot configurations can be easily configured via a graphical user-interface as well as non-interactively via a script. The large variety of plot options renders rfxplot suitable both for data exploration as well as producing high-quality figures for publications.

Introduction

Statistical Parametric Mapping (SPM) is a widely used software package for analyzing neuroimaging data (http://www.fil.ion.ucl.ac.uk/spm) which utilizes a mass univariate approach for determining the parameter estimates of a general linear model (GLM) at each voxel in the brain (Kiebel and Holmes 2007). These parameter estimates (or effect sizes) can be then subjected to a statistical test (contrast), which results in a statistical parametric map (Poline et al. 2007). These statistical maps can accurately display the location and spatial extent of an experimental effect. However, they fall short in conveying the "nature" of more complex effects like interactions in a factorial design (e.g., which condition in the interaction effect yields the largest activation). Thus, there exists an obvious necessity for plotting the underlying data, especially for the more complex multi-factorial designs that are widely employed in contemporary neuroimaging studies.

The analysis of functional neuroimaging data in SPM proceeds in two stages: at the first level, subject-specific GLMs, which include regressors of all expected experimental effects and additional nuisance variables, are estimated and parameter estimates are derived. Building on the summary statistics approach (Penny and Holmes 2007), these individual parameter estimates are then subjected to a second level analysis for a statistical test across the entire sample, thus allowing for population

J. Gläscher (✉)
Division of Humanities and Social Sciences,
California Institute of Technology,
1200 E California Boulevard,
Pasadena, CA 91125, USA
e-mail: glascher@hss.caltech.edu

inference. Early functional neuroimaging in the 1990s combined all first level analyses into one single analysis, thus limiting the inference to merely the actual sample that was scanned. Nowadays, however, statistical inference is drawn from second level analyses that treat subjects as a random effect drawn from a larger population. Because individual subjects are regarded as a random effect, these second level analyses are commonly referred to as "random effects analyses".

SPM offers a variety of different options for plotting data from first level analyses (e.g., contrast estimate, fitted responses, event-related responses), but these options are quite sparse for second level analyses. rfxplot is a new toolbox designed to alleviate this problem. The toolbox provides an integrated and highly configurable approach for data visualization from within second level analyses. The name "rfxplot" was chosen because it conveys that the starting point of the toolbox is a second level random effects (RFX) analysis. rfxplot is distributed with an extensive user manual and can be downloaded at http://rfxplot.sourceforge.net.

Overview of the Toolbox

The toolbox fully integrates into SPM and uses its windows for entering configuration information and displaying data plots (Fig. 1a). In addition, the data can be also displayed in other MATLAB figure windows making it easy to assemble several panels within a larger figure for a publication. User-specified options for plot configuration are entered via buttons, text entry fields, and drop-down menus in rfxplot input area (see Fig. 1 for examples of the input interface).

Plot Selection

Three types of data are most commonly visualized in publications from neuroimaging studies: (1) average parameter estimates across subjects extracted from the individual first level analysis, which can be interpreted as the average effect size or activation of all trials that are modeled in the corresponding regressors, (2) fitted responses, which convolve the parameter estimate with the basis function to visualize the response shapes, and (3) event-related BOLD responses (also known as BOLD time courses or peri-stimulus time histograms, PSTHs), which show the average evoked BOLD response of all trials belonging to an experimental condition analogous to an event-related potential (ERP) recorded with scalp EEG recordings.

rfxplot currently offers all of the above mentioned data plots, which can be selected in a drop-down menu (Fig. 1b): (1) Plots of parameter estimates (effect sizes) are the most efficient ways to visualize a particular

experimental effect, especially for multi-factorial designs. They are particularly suited for experiments in which the data are modeled with a single canonical hemodynamic response function (HRF), because the parameter estimate scales the height of the HRF and can therefore be directly interpreted as the average BOLD activation for a particular experimental condition. With the proper scaling, these average parameter estimates can be expressed as percent signal change (see below). (2) Plots of fitted responses are more suitable in situation in which the BOLD response is modeled with a high-dimensional set of basis function (e.g. the canonical HRF and its temporal and dispersion derivatives, or with a Fourier basis set). These plots can reveal different response shapes between conditions (e.g. early vs. late BOLD responses). (3) Event-related BOLD responses display the average height of the BOLD responses in a user-defined time window before and after the stimulus onset. This plot shows the actual data and does not rely on any first level model. Event-related BOLD responses are estimated by a condition-specific Finite Impulse Response (FIR) model. The post-stimulus time in which the BOLD response is expected to rise is parcellated into different temporal bins starting at or shortly before the onsets of all trials belonging to a particular condition. Different regressors are created for each bin with 1's at each data point belonging to this bin, and 0's elsewhere (see Fig. 2 for an example). No intercept term is included and the parameters of the FIR model are estimated using standard ordinary least squares. The parameter estimate for each bin of this FIR model is identical to the mean BOLD response in that bin, thus creating an event-related BOLD time course. The toolbox takes special care to preserve the integrity of the FIR model when the length of the time window is too long for closely spaced trials (see Fig. 2 for details on the procedure). For the final display, rfxplot averages these subject-specific event-related BOLD responses.

Voxel Selection

The selection of voxels from which the data are extracted and plotted is an important decision that can affect the visual display dramatically. Possible selection strategies can be loosely grouped on a single dimension ranging from a strict selection of significant voxels in an RFX analysis to a subject-specific selection of voxels that are based on the maximum effect in an individual activation pattern.

Whereas the first selection strategy offers the greatest anatomical precision (in the extreme case only a single voxel), it runs into the danger of missing stronger activations nearby in some subjects. It also disregards that considerable variability in the individual location of brain regions remains even after stereotactic normalization. Therefore, the selected voxels from an RFX analyses may

Fig. 1 Examples of the graphical user interface of rfxplot. **a** rfxplot uses existing SPM windows for entering plot configurations (input area) and for displaying the data (display area). **b** Several different plot types (effect sizes, fitted responses, event-related BOLD responses) can be selected via a drop-down menu. **c** Voxel selection proceeds in two steps. First, an RFX search volume is defined via a drop-down menu. Later, the selection of voxels from individual first level analyses can be refined. **d** Images are selected by references to a table of image descriptions. Provided these descriptions are consistent across subjects, rfxplot can deal with missing images in some of the subjects. **e** Regressors can be split into different bins, thus allowing for a visualization of only certain trials of a single regressor. These bins are configured by a series of text inputs
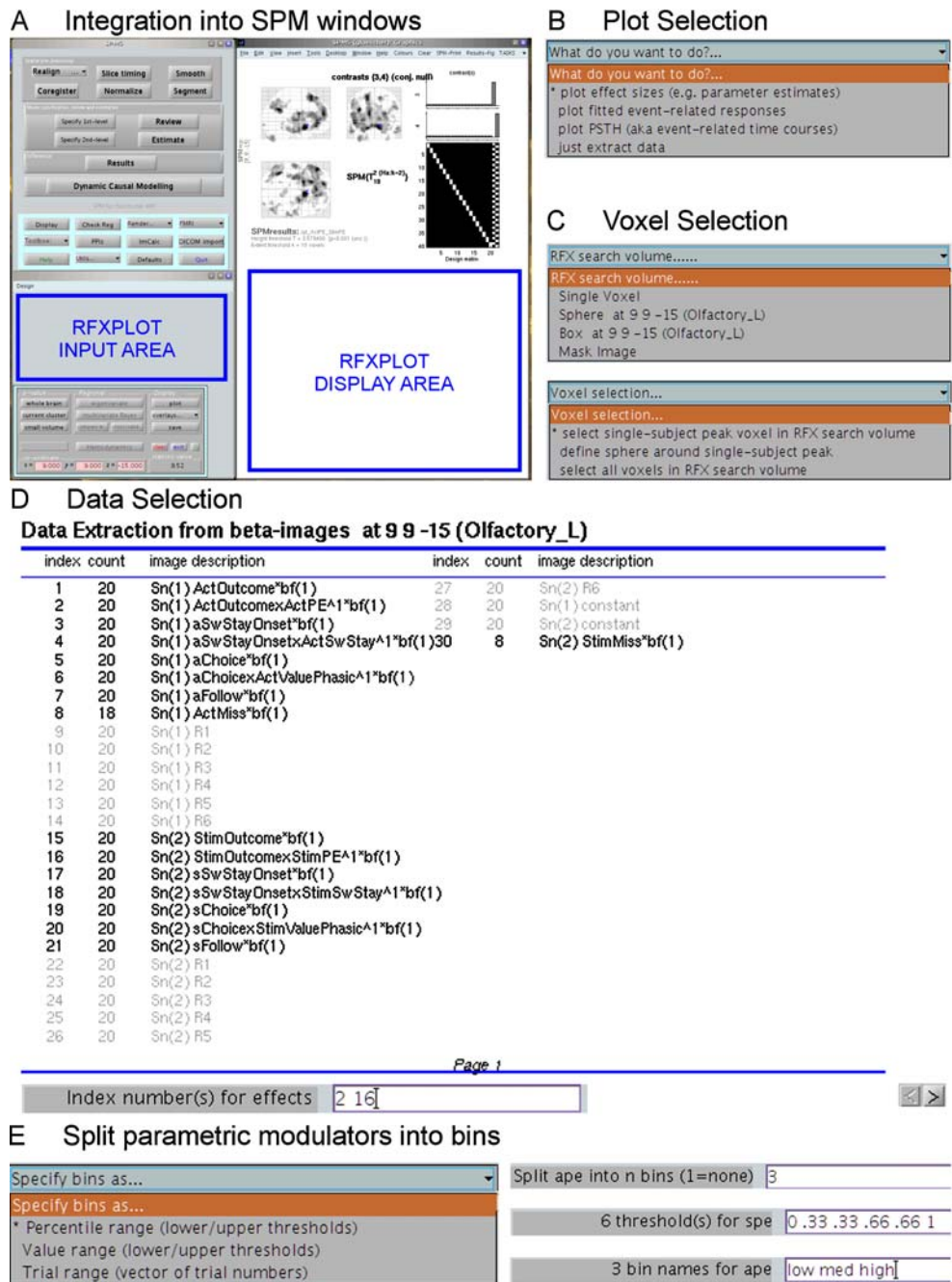
not fall into the identical anatomical region in all subjects. On the other hand, a selection of significant voxels from the individual first level analysis can result in a data plot that, while maximizing the individual effect size, is not representative of any common anatomical region for all subjects.

rfxplot covers the middle ground between these extremes and offers a flexible two-step voxel selection process. First, a search volume based on an RFX contrast is defined with respect to the current voxel (single voxel, sphere, or box), or independently by selecting an image mask defining an anatomical region (Fig. 1c upper drop-down menu). It is

also possible to select subject-specific image masks, which delineate a certain region in every subject. In addition, the RFX search volume can comprise *all* voxels or it can be limited to include only the significant (supra-threshold) voxels in the current second level contrast.

In a second step, the selection of voxels can be refined by limiting it to the voxel with the maximum effect size in each subject given the current second level contrast (Fig. 1c lower drop-down menu). Thus, it possible to account for different locations of the maximum effect across subjects, while limiting the search volume in where to look for these peaks to a specific anatomical area. The toolbox determines
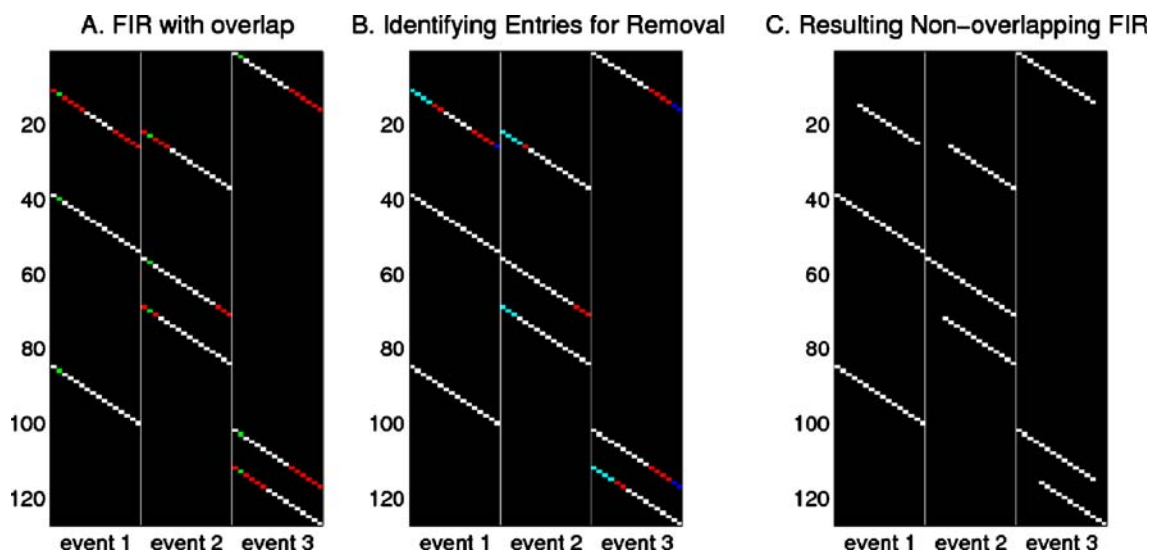
Fig. 2 Example for estimating event-related BOLD responses with a Finite Impulse Response (FIR) model. **a** Original FIR design for three different event types with three trials each. The FIR for each event contains as many regressors as specified time bins (16 in this example). The stimulus onset is marked by the *green* bin. Due to the relatively long time window and the rapid presentation time, some of the bins are overlapping between and within event types (marked in *red*). Estimating this FIR design would yield parameter estimate for these bins that are not identical to the mean BOLD signal in each bin. **b** rfxplot resolves this ambiguity by removing these entries from the FIR design. Assuming that the most important part of the BOLD response is its peak approximately 4–5 s after stimulus onset, rfxplot starts removing entries from the beginning of the time window until 2 s after stimulus onset (marked in *cyan*). If this does not completely resolve the overlap, then overlapping entries are removed from the end of the time window (marked in *blue*). **c** The resulting FIR design after the removal of overlapping bins. This procedure implemented in rfxplot tries to preserves the bin structure with an emphasis on those bins when the maximum BOLD response is expected (~4–5 s after stimulus onset). However, it also insures that the parameter estimates for the FIR design will always be identical to the average BOLD signal in each bin

the maximum effect size by retrieving the data from all images of each subject in the search volume and multiplying them with the current second level contrast weights, i.e., max effect=max($c^T y$), for each subject, where c is the current second level contrast, and y are the (second level) data for each subject. This approach insures that the maximum effect for each subject can be computed for all second level contrasts (simple and differential effects).

Optionally, an additional spherical search volume can be defined around the individual peaks in order to retrieve the data from a more representative set of voxels than just the peak in each subject (Fig. 1c lower drop-down menu). In practice, however, this step will often prove to be unnecessary, because the data are spatially smoothed during image preprocessing and hence the effect sizes in these individual spheres will have a similar magnitude as the peak voxel itself.

Data Selection

The selection of regressors from which to extract the data is equally important as the selection of voxels. While a certain effect is chosen for an RFX contrast and is then used for voxel selection, it might be also informative to display the data from different regressors in order to obtain a complete impression of the underlying activation pattern in a

particular brain region. These comprehensive data plots can protect against the potentially false inference of claiming exclusive regional specificity for a single stimulus class.

The toolbox addresses this issue by providing a comprehensive set of images to choose from for displaying the underlying data. In fact, the user can choose to plot the data from any regressor that has been included in the design matrix or from any contrast that has been specified. For instance, given a statistical map based on a functional localizer for the fusiform face area (FFA) it is possible to visualize the data in the FFA not only for activations to faces, but also to other stimuli (like objects).

The toolbox will always retrieve the data from the underlying first level analyses. This avoids some possible "rescaling" issues that arise in second level SPM analyses due to the ubiquitous over-parameterized design matrices in which subject-specific and/or group-specific intercept terms are included and which model the mean effect of a subject or group. As a consequence, the parameter estimates for the other regressors (of the experimental conditions) only reflect the difference from this mean effect size. While this procedure results in parameter estimates that are more easily interpretable and can increase statistical power, it also has the drawback that the original scale of the first level estimates is lost and "activations" cannot be discriminated from "deactivations". rfxplot avoids these problems

by always retrieving the data from the first level analyses and averaging them directly.

Finally, rfxplot selects the images for retrieving the individual data by the image description (Fig. 1d). This selection procedure protects against different first level design matrices across subjects provided that the conditions are named consistently. For example, in a study on the subsequent memory effect, the first level regressors are divided by the subject's post-experimental memory performance (Staresina and Davachi 2006). Trials that were later remembered are collected in one regressor, whereas trials that were later forgotten are encoded in a different regressor. This can result in different first level design matrices because due the subject's performance one of the two regressors may be missing. However, the toolbox can deal with this situation if the different conditions ("remembered", "forgotten") are named consistently across subjects.

Splitting Regressors into Different Bins

In addition to "onset regressors", which model a canonical BOLD response to a given stimulus at a particular point of time in the experiment, SPM also provides a way to include a second regressor (called a "parametric modulator"), which can modulate the height of the modeled BOLD response on a trial-by-trial basis. The parameter estimates of these modulators signify how well the BOLD signal in a certain area correlates with a more complex and continuous variable. For instance, by including trial-specific reaction times (RTs) as a parametric modulation, one can detect brain areas that covary with RTs, possibly indicating brain regions that respond to higher processing demands. In the above mentioned subsequent memory study (Staresina and Davachi 2006), it would be equally possible to encode memory success as a parametric modulator rather than using two separate regressors for memory success or failure. These modulators can be also used in the context of applying computational models of learning to neuro-imaging data. For instance, model-derived prediction errors can be used as modulators, thus allowing the detection of brain regions that encode a prediction error signal (O'Doherty et al. 2003, 2004).

These parametric modulators are orthogonalized with respect to the onset regressor, but are subsequently treated as another regressor in the GLM. As a consequence, the parameter estimate (beta) of the onset regressor will reflect the height of the average activation to all trials in a particular condition, whereas the beta of the modulator will express how well the residual fluctuations in the BOLD signal covary with the modulator. Consistent with its interpretation as a regression coefficient, a positive beta for the modulator indicates a higher BOLD response for trials with a high parametric value and a smaller BOLD response for trials with a low parametric value. However, this interpretation is rather abstract and a more direct visualization of this relationship usually facilitates comprehension.

rfxplot is designed to provide this kind of data plot. The events of a particular regressor (parametric modulators and onset regressors) can be split into different bins according to three criteria: (a) according to a percentile range, (b) according to the absolute values of a modulator, or (c) according to trial numbers (Fig. 1e left drop-down menu). For instance, in the first case (percentile range) one could divide the modulator into three equally sized bins (low, medium, high) according to the 33rd, 66th, 100th percentile (Fig. 1e right text input fields). In the second case (absolute values), one could split a modulator encoding RTs into two bins, one containing all trials with RTs shorter than 250 ms and another one containing all trials with RTs longer than 250 ms. Finally, both modulators and onset regressors can be divided according to trial number (e.g. splitting a regressor into the first and second half of the experiment). rfxplot will re-sort all trials in a regressor according to these newly created bins, create a new onset regressor for each bin, and estimate a beta signifying the average height of the BOLD response for all events in each bin.

The advantage of using a percentile range for defining the different bins is that it is independent of the range of parametric values in each subject. It also ensures that there are approximately an equal number of events in each bin. Defining the bins by absolute values is possible when the range of the parametric variable is clearly defined or when there is a special interest in parametric values above or below a certain threshold (e.g. a RT of 250 ms). Finally, splitting a regressor according to trial numbers provides an easy way of assessing possible effects of time (e.g. that the BOLD response is smaller in the second half of the experiment indicating potential habituation effects). The splitting of regressors into different bins is compatible with all plot types mentioned in the section "Plot Selection".

Splitting the Sample into Different Groups

The toolbox is mainly targeted for data inspection of RFX analyses and therefore provides different options for assessing the between-subject variability in a group of subjects. A possible reason for large variability might be that there are unrecognized groups within the sample of subjects that respond differently to the experimental paradigm. When modeled as a single group of subjects these effects can increase the unexplained variance and reduce statistical power. Therefore, an easy way of inspecting possible effects of sub-groups would be beneficial for more detailed future analyses.

The toolbox provides the option of dividing the subjects into different groups (e.g. male–female, young–old, learners–non-learners) suitable for a quick inspection of potential underlying group differences that might cause a different BOLD response profile. If effects of unrecognized subgroups are apparent in a data plot, then this can serve as starting points for more detailed analyses of each of these subgroups. Furthermore, the individual data points for each subject (and their subject numbers) can be overlaid onto the plot. This can facilitate the detection of outlier subjects which can be excluded in a subsequent plot or analysis.

Percent Signal Change

Over the past few years, Percent Signal Change (PSC) has become a de facto standard for reporting the magnitude of activations in functional neuroimaging studies. It is an unfortunate and regrettably common practice to assume that the second level parameter estimates (betas) are expressed in PSC because the EPI images are rescaled to a session-wise mean of 100 during SPM image preprocessing. There are two reasons why this is not appropriate. First, the beta of a canonical HRF scales the height of the regressor, thus transforming it into a fitted response. Therefore, it is actually the height of the fitted response that should be taken as an estimate of PSC. Secondly, because the scaling during image preprocessing applies to the entire image the scale of the betas can be at best called Percent *Global* Signal Change because the scaling does not take the differences between different tissue types into account.

rfxplot alleviates these two problems and computes PSC by taking the mean signal intensity of the tissue type into account which is captured in the voxel-specific and session-wide constant term in a first level analysis. rfxplot computes PSC with the following equation:

$$PSC = \frac{\beta_{\text{task}} \times \max(HRF) \times 100}{\beta_{\text{const}}}$$

where β_{task} is the parameter estimate of the condition of interest in the voxels of interest, $\max(HRF)$ is the maximum of a regressor with a single event, and β_{const} is the parameter estimate of the session-wide constant for the particular voxels of interest. In that sense, the PSC computed by rfxplot can be regarded as a measure of Percent *Local* Signal Change which expresses how much a particular voxel is activated by an experimental conditions compared to its own baseline.

Fitting a Regression Model to Extracted Effect Sizes

A whole family of neuroimaging experiments is devoted to investigating the relationship between external stimulus properties and the magnitude of the BOLD response in areas thought to be dedicated to processing the stimulus. One of the first examples was the relationship between image contrast and activation of V1 (Boynton et al. 1996), which demonstrated larger BOLD responses with increasing image contrast. When plotting the effect size for different degrees of image contrast with rfxplot, the toolbox also offers to fit a regression model to these data. This can be used to test for a linear increase in V1 activation when image contrast increases. rfxplot will fit this secondary regression model to both the average data across subjects, as well as to the data of each individual subject. While the former is displayed in the data plot of the average activation across subject, the latter can be used for statistical tests of model parameters (e.g. a one-sample *t*-test of the individual slope parameters in the linear fit). Currently, rfxplot supports four regression models: linear, quadratic, exponential, and logarithmic.

Features for Re-using Plot Configurations

Data plots can be configured via a graphical user interface that is integrated into SPM (Fig. 1a). The plot configuration dialogue consists of a series of questions requiring user input. Because this dialogue is quite long, rfxplot provides several ways for circumventing it. First, if rfxplot finds a plot configuration in the current MATLAB workspace it will offer the options of re-using this previous configuration. This option is mostly useful for creating the same data plot at different voxels in the brain. Furthermore, each configuration can be saved under a specific name and is available for re-usage in a later SPM session. Finally, as detailed in the user manual, plot configuration can be scripted, thus allowing a completely non-interactive usage of rfxplot.

Results

This section provides some examples of the data plots created with rfxplot. The data used for these plots are taken from Glascher et al. (2007, 2008).

In the first study (Glascher et al. 2008) we investigate the neural correlates of value and prediction error signals derived from a reinforcement learning model in a probabilistic reversal learning task. A more detailed description of task, methods, and findings can be found in the original publication. Both signals were modeled as parametric modulators in the subject-specific first level analyses. Parameter estimates for value and prediction error signals were subsequently tested at the second level in a one-sample t-test across subjects. Strong effects for the prediction error signals were found bilaterally in the ventral striatum. Figure 3 shows different variants of data plots of
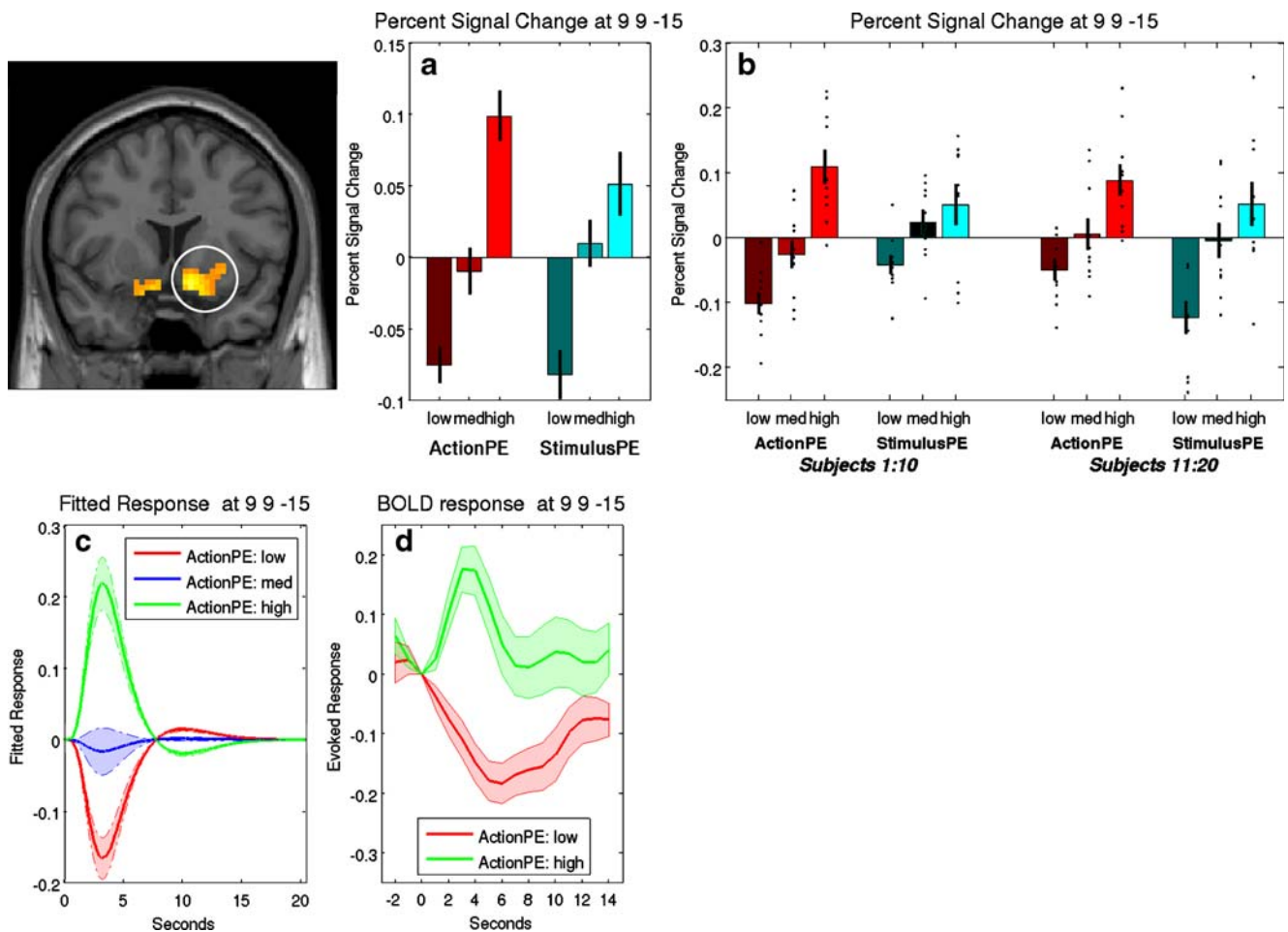
Fig. 3 Examples of data plots created with rfxplot (data from Glascher et al. 2008). All *error bars* are s.e.m. across subjects. **a** Mean percent signal change for a parametric modulator encoding a model-derived prediction error signal in the ventral striatum, which has been split into three equally sized bins. **b** Same as in (**a**) but with subjects split into two arbitrary groups. Individual data points are superimposed on the bar graphs. **c** Fitted responses for the three different bins of parametric modulator in the action reversal condition. **d** Event-related BOLD responses for the trials with low vs. high prediction errors. Low=bottom quartile, high=top quartile of possible range of prediction errors

the prediction error signal created with rfxplot. Voxels were selected by placing an 8 mm spherical search volume around the peak coordinate in the RFX analysis (x, y, z=9, 9, −15) and subsequently selecting the peak voxel for each subject within that search volume. Finally, a smaller 6 mm sphere was place around the individual peak voxels and data were extracted from this volume and averaged.

Figure 3a shows a plot of the effect size in the two experimental conditions (action and stimulus reversal) when the parametric modulator (model-derived prediction error) was split into three equally sized bins according to the 33rd, 66th and 100th percentile. Colors, error bars (s.e.m. across subjects), and axis labels were all configured in rfxplot. Figure 3b shows the same data (error bars: s.e.m. across subjects), but now the sample is split into two arbitrary groups of subjects with the individual data point superimposed onto each bar representing the average percent signal change for a particular bin (low, medium,

high prediction errors). Figure 3c shows the same data for the action reversal condition but in the format of fitted responses (error bars: s.e.m. across subjects). Finally, Figure 3d displays the same data for the action reversal condition for only the bottom and top quartile of possible prediction errors (error bars: s.e.m. across subjects). These quartiles were binned by percentile ranges with the low bin covering the 0th–25th percentile and the high bin covering the 75th–100th percentile. In contrast to the plot in 3c these event-related BOLD response reveal that the (positive) response to trials with a high prediction error peaks earlier than the (negative) peak to the trials with low prediction error. This information could not have been detected by the plot of fitted response because this study only used the canonical HRF as a basis function, which models the response rigidly at the same point in time.

The second example utilizes the data from Glascher et al. (2007). This study investigated the influences of

cognitive load and emotionality of the stimulus on the activation in extra-striate visual areas. The experiment was a three-factorial design with the factors working memory load (1-back or 2-back task), emotion (neutral or negative background stimuli), and visibility (levels 0%, 33%, 66%, and 100%). The latter factor was realized by scrambling the phase information of the background stimuli. Stimuli were presented in blocks of 10 pictures and subjects had to perform either a 1-back or a 2-back working memory task on a color square presented in the image foreground. The details of the experimental protocol and results can be found in the original publication.

Figure 4a shows the fitted responses (error bars: s.e.m. across subjects) as semi-transparent "error regions") for

all four visibility conditions in the lateral occipital complex which showed a strong main effect of visibility (i.e. increase in activation with higher levels of visibility). Figure 4b shows the parameter estimates of the conjunction of the two-way interactions of emotion × visibility and load × visibility. The average parameter estimate across subjects is represented by a dot with black error bars (s.e.m. across subjects). Linear regression models were fitted to these average estimates indicating a significant increase in activation across the levels of visibility. The slopes for the two different n-back conditions (1-back in red, 2-back in blue) were significantly different, as tested by a paired t-test on the slope parameters from the linear regression for each subject.
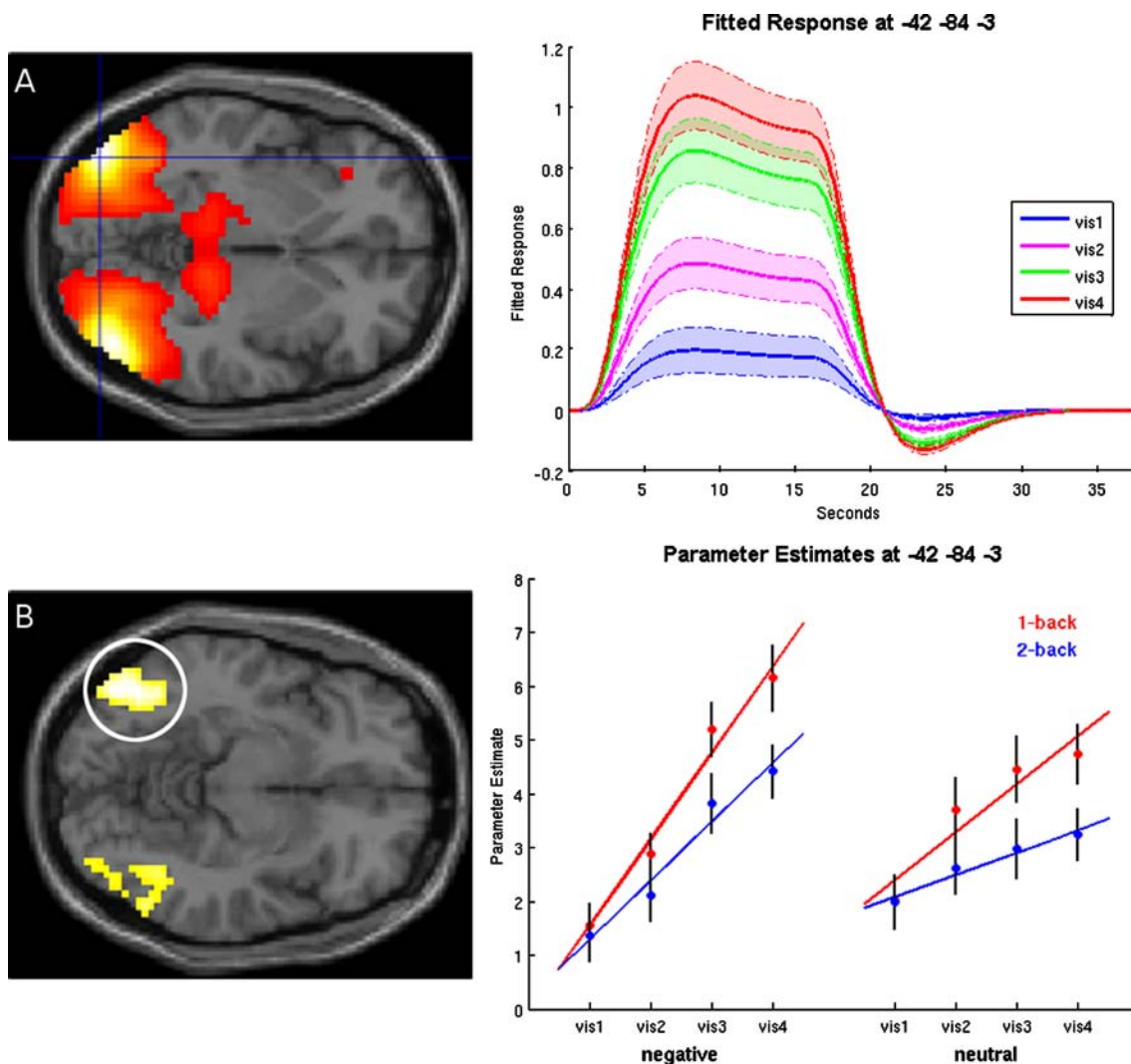

Fig. 4 Example of data plots created with rfxplot (data from Gläscher et al. 2007). All *error bars* are s.e.m. across subjects. **a** Fitted responses for the different experimental levels of the main effect of visibility. **b** Parameter estimates for the conjunction of the two-way interaction of emotion × visibility and load × visibility. The regression lines were fitted with rfxplot to the average data for both levels of working memory load (1-back, 2-back) across the different levels of visibility

Comparison with Existing SPM Toolboxes

The voxel selection process of rfxplot combines approaches from whole-brain, voxel-wise statistical tests with those typically found for Region of Interest (ROI) analyses. The latter are implemented in the MarsBar Toolbox (Brett et al. 2001). In contrast to whole-brain voxel-wise statistical analyses, the time series of all voxels with an ROI are collapsed into a single time series, which is then analyzed using standard GLMs. ROI analyses therefore allow for voxel selection based on the individual activation profile, thus tailoring it to the maximum effect size. As explained in detail above, the voxel selection procedure of the toolbox can be seen as the middle ground between these two approaches. Unlike the implementation of ROI analyses in MarsBar, however, rfxplot is a tool for data visualization and will not run any statistical analyses on the data it extracts and displays.

The estimation of event-related BOLD responses via an FIR model is akin to existing approaches implemented in SPM itself and in an additional tool developed by Darren Gitelman (get_psth.m, downloadable at http://brainimaging.tiddlyspot.com/). However, unlike these previous implementations rfxplot offers a larger flexibility in selecting regressors and splitting a regressor into different bins, thus allowing for different event-related BOLD responses *within* a single regressor.

Conclusion

rfxplot provides a variety of options for creating data plots of neuroimaging data (effect size, fitted response, and event-related BOLD responses) from within RFX analyses in SPM. This approach is congruent with the scientific interest of the neuroimaging community in group level analyses, which allow for population inference. Nevertheless, the toolbox also offers various features for examining individual subjects (e.g. by overlaying subject-specific data points) and, to a certain extent, the microstructure of trials (e.g. by splitting a regressor into a number of different bins). These visualization techniques can therefore help to detect inter-individual differences in functional neuroimaging studies, an area that is becoming increasingly important in the field. Many experimenters nowadays realize that the large extent of between-subject variability is not only a risk to statistical power that needs to be overcome by increasing sample sizes, but rather that interesting and meaningful information is contained in this inter-subject variability that can be explained by additional external variables. In that sense, many researchers are seeking out additional per-subject variables that can explain parts of this variance and thus increase statistical power and characterize the data in greater detail. A recent trend in several publications is to

correlate the size of activation in a particular brain area with external variables such as measures derived from personality questionnaires (Kumari et al. 2004) or even with poly-morphic variants in a subject's genetic profile (Yacubian et al. 2007), as it is frequently done in the emerging field of imaging genomics (Meyer-Lindenberg and Weinberger 2006). To further support this goal of examining between-subject variability, it is intended that a future version of the toolbox will be able to create correlation plots of BOLD activation data with any arbitrary external variable and to fit different regression models to these data (e.g. linear, curvilinear, robust regression). These enhancements, together with the already existing highly configurable visualization features and the user-friendly interface, make rfxplot ideal for data exploration and creating high-quality figures for publications.

References

Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *The Journal of Neuroscience*, 16, 4207–4221.

Brett, M., Anton, J-L., Valabregue, R., & Poline, J-B. (2001). Region of interest analysis using an SPM toolbox [abstract] Presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2-6, 2002, Sendai, Japan. Avaible on CD-ROM in NeoroImage, Vol 16, No 2.

Glascher, J., Rose, M., & Buchel, C. (2007). Independent effects of emotion and working memory load on visual activation in the lateral occipital complex. *The Journal of Neuroscience*, 27, 4366–4373. doi:10.1523/JNEUROSCI.3310-06.2007.

Glascher, J., Hampton, A. N., & O'Doherty, J. P. (2008). Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. June 11, 2008. doi:10.1093/cercor/bhn098.

Kiebel, S. J., & Holmes, A. J. (2007). The general linear model. In R. S. Frackowiak, K. J. Friston, C. D. Frith, R. J. Dolan, C. J. Price, S. Zeki, J. Ashburner, & W. D. Penny (Eds.), *Human brain function* (pp. 101–126). San Diego: Elsevier.

Kumari, V., ffytche, D. H., Williams, S. C., & Gray, J. A. (2004). Personality predicts brain responses to cognitive demands. *The Journal of Neuroscience*, 24, 10636–10641. doi:10.1523/JNEUROSCI.3206-04.2004.

Meyer-Lindenberg, A., & Weinberger, D. R. (2006). Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nature Reviews. Neuroscience*, 7, 818–827. doi:10.1038/nrn1993.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, 38, 329–337. doi:10.1016/S0896-6273(03)00169-7.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, *304*, 452–454. doi:10.1126/science.1094285.

Penny, W. D., & Holmes, A. J. (2007). Random-effects analysis. In R. S. Frackowiak, K. J. Friston, C. D. Frith, R. J. Dolan, C. J. Price, S. Zeki, J. Ashburner, & W. D. Penny (Eds.), *Human brain function* (pp. 156–165). San Diego: Elsevier.

Poline, J. B., Kherif, F., & Penny, W. D. (2007). Contrasts and classical inference. In R. S. Frackowiak, K. J. Friston, C. D. Frith, R. J. Dolan, C. J. Price, S. Zeki, J. Ashburner, & W. D.

Penny (Eds.), *Human brain function* (pp. 126–140). San Diego: Elsevier.

Staresina, B. P., & Davachi, L. (2006). Differential encoding mechanisms for subsequent associative recognition and free recall. *The Journal of Neuroscience*, *26*, 9162–9172. doi:10.1523/JNEUROSCI.2877-06.2006.

Yacubian, J., Sommer, T., Schroeder, K., Glascher, J., Kalisch, R., Leuenberger, B., Braus, D. F., & Buchel, C. (2007). Gene–gene interaction associated with neural reward sensitivity. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 8125–8130. doi:10.1073/pnas.0702029104.