Deconvolution of Event-Related fMRI Responses in Fast-Rate Experimental Designs: Tracking Amplitude Variations

Hermann Hinrichs, Michael Scholz, and Claus Tempelmann
University of Magdeburg, Germany

Martin G. Woldorff
University of Magdeburg, Germany and Duke University

Anders M. Dale
NMR Center, Massachusetts General Hospital

Hans-Jochen Heinze
University of Magdeburg, Germany

Abstract

■ Recent developments towards event-related functional magnetic resonance imaging has greatly extended the range of experimental designs. If the events occur in rapid succession, the corresponding time-locked responses overlap significantly and need to be deconvolved in order to separate the contributions of different events. Here we present a deconvolution approach, which is especially aimed at the analysis of fMRI data where sequence- or context-related responses are expected. For this purpose, we make the assumption of a hemodynamic response function (HDR) with constant yet not predefined shape but with possibly variable amplitudes. This approach reduces the number of variables to be estimated but still keeps the solutions flexible with respect to the shape. Consequently, statistical efficiency is improved. Temporal variations of the HDR strength are directly indicated by the amplitudes derived by the algorithm. Both the estimation efficiency and statistical inference are further supported by an improved estimation of the noise covariance. Using synthesized data sets, both differently shaped HDRs and varying amplitude factors were correctly identified. The gain in statistical sensitivity led to improved ratios of false- and true-positive detection rates for synthetic activations in these data. In an event-related fMRI experiment with a human subject, different HDR amplitudes could be derived corresponding to stimulation at different visual stimulus contrasts. Finally, in a visual spatial attention experiment we obtained different fMRI response amplitudes depending on the sequences of attention conditions. ■

INTRODUCTION

Event-related functional magnetic resonance imaging (ErfMRI) (Rosen, Buckner, & Dale, 1998; Dale & Buckner, 1997; Josephs, Turner, & Friston, 1997) has greatly extended the range of questions that can be addressed by functional MRI. In contrast to the traditional block-design fMRI experiments, with ErfMRI fully randomized orders of events may be used. This not only facilitates the integration of fMRI with electrophysiological measurements, such as those using event-related potentials (ERPs), but also avoids potential confounds like habituation or strategy effects.

Advanced fMRI studies usually apply interstimulus intervals (ISI) in the range of a few seconds or below. Taking into account that the hemodynamic response (HDR) elicited by an event extends over 10–15 sec (Boynton, Engel, Glover, & Heeger, 1996) subsequent HDRs will overlap significantly in this type of experiments. Therefore, simple event-related averaging often will not be sufficient to extract the ErfMRI response. First ideas to get rid of this overlap were proposed by Hansen (1983), Woldorff (1993), and Ganis, Kutas, Schendan, and Dale (1997) for the analysis of ERPs.

A general framework for estimating evoked HDRs using the general linear model (GLM) or multiple linear regression was outlined in Josephs et al. (1997). In this paper, we propose the use of basis functions with continuous support that allowed for distributed sampling of the data over peri-stimulus time. A special case

of this analysis obtains when the basis set corresponds to nonoverlapping time bins that correspond to discrete sampling of points following stimulus onset. In this instance, the basis function approach reduces to a finite impulse response function estimation. This approach has been subsequently proposed by Dale (1999) and Burrock and Dale (in press) and is a useful parameterization in the sense that the parameter estimates conform to the underlying HDR function. The number of basis functions or number of time bins considered can sometimes lead to overfitting of the model or inefficient parameter estimates. In this instance, constraints are required. Josephs et al. (1997) and Friston et al. (1998) have proposed the use of a small set of basis functions (two or three). The parameter estimates or coefficients of these basis functions are unbiased estimators of the underlying HDR projected onto the subspace of the basis set. However, unless the assumed basis functions span the entire space of possible response functions, the resulting response function estimates may be systematically distorted.

In this paper, we take a similar approach but enforce more natural constraints that may be useful in specific experimental situations. Put simply, we constrain the shape of the HDR to be the same for a particular class of stimuli or tasks but allow its amplitude to vary among different trial types within that class. To implement this constraint we have to estimate the amplitudes of the different trial types and a common shape for the HDR. This parameter estimation problem precludes a conventional least squares approach and requires an iterative parameter reestimation scheme that is described below.

In the following we first give a brief overview over the GLM approach according to Dale (1999) and Burock and Dale (in press) giving the basic formula both for the HDR estimate and for the corresponding errors. The original equations will be modified by introducing the weighting factors (WFs) in the context of the general equations. From these equations it follows immediately that estimating the HDR and the set of WFs at the same time leads to nonlinear equations that cannot be solved directly. We therefore developed a nonlinear steepest descent procedure that iteratively optimizes the HDR and the set of WFs with the goal of minimizing the residual error. We will also describe an improvement to the Burock and Dale procedure to estimate the temporal correlations of the raw fMRI data needed to set up the covariance matrix, which is part of the GLM model.

In order to check the new algorithm we performed simulations on synthesized data sets. The results prove both the general feasibility and the statistical efficiency of the new algorithm. Finally, two experimental examples are presented with the goal of demonstrating practical applications of the new method. They also should give an idea of how to decide whether the new procedure may be applied to a certain experiment.

MATHEMATICAL BACKGROUND

Detrending of MRI Data

Measurement related systematic shifts ('trends') are removed by an automatic trend removal procedure as follows:

In each run (i.e., each time period of continual MR data acquisition), comprising Nr measurements, a window of $Nr/2$ samples is shifted over the Nr samples, starting and ending with the window centered around the first and last sample respectively. For the first and last $Nr/4$ samples the window is truncated, resulting in a few less samples per window in this situation. For each of the windows over the run, a modified median value is calculated as the mean of the 45- and 55-percentile values. The median is subtracted from the original sample yielding the trend-corrected value. As compared to an ordinary polynomial regression the procedure presented here is more robust against outliers and more flexible with respect to the temporal characteristic of the trend. Examples are given in Figure 1.

General Framework of the Linear Model Allowing for Amplitude Variations

Assuming a linear time invariant model as suggested by Boynton et al. (1996), Dale and Buckner, (1997), Dale (1999) and others, the fMRI signal $y(t)$ observed during stimulation is given by

$$y(t) = d(t)*h(t) + n(t) \tag{1}$$

where $d(t)$ is a sequence of time-shifted delta functions indicating the onset of the subsequent stimuli, $n(t)$ is a Gaussian process specifying the measurement noise, and $h(t)$ is the underlying HDR. For experiments with different event types this equation may be generalized to

$$y(t) = \sum_{i=1}^{N_c} d_i(t)*h_i(t) + n(t) \tag{2}$$

with N_c specifying the number of different event types and $d_i(t)$ and $h_i(t)$ representing the pulse train and HDR respectively for event type i. In order to allow for individual HDR amplitudes the pulse trains $d_i(t)$ may be replaced by modified versions with nonuniform WFs, i.e.,

$$y(t) = \sum_{i=1}^{N_c} x_i(t)*h_i(t) + n(t) \tag{3}$$

where $x_i(t)$ is defined as

$$x_i(t) = \sum_{j=1}^{Nw_i} w_{i,j} e_{i,j}(t) \tag{4}$$

with the constraint

$$d_i(t) = \sum_{j=1}^{Nw_i} e_{i,j}(t) \tag{4a}$$

According to Equation (4) the pulse sequences of each event type i is subdivided into Nw_i groups of

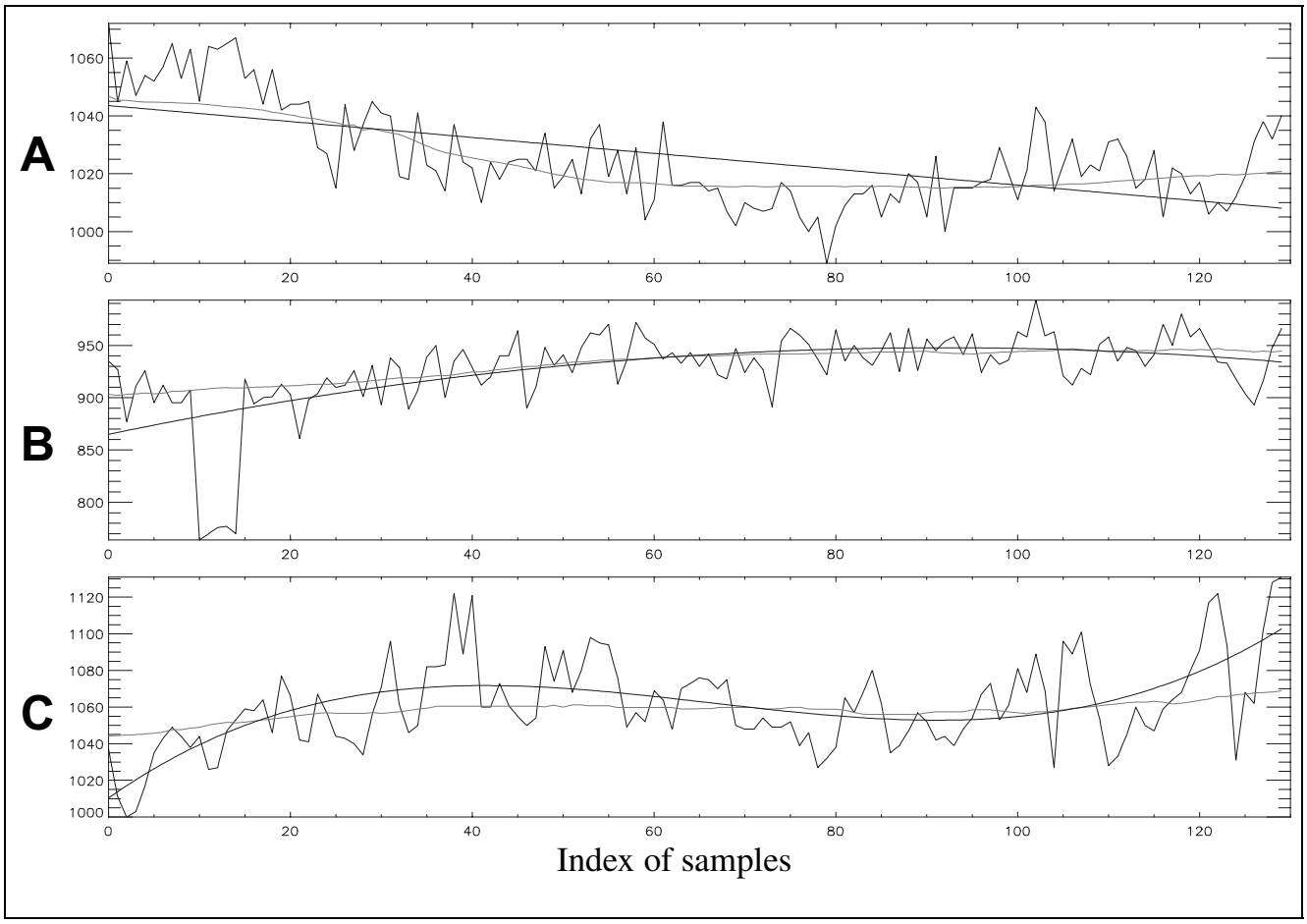

Figure 1. Trend removal in raw fMRI data. Comparison of trends derived by median method (gray line) and first-order (A), second-order (B) and third-order (C) polynomial (black line). The example shown in trace B was chosen in order to show that the trend removal is robust with respect to outliers.

events with individual weights $w_{i,j}$ and a corresponding pulse train $e_{i,j}(t)$ indicating the events where $w_{i,j}$ applies. Equation (2) is now a special case of Equations (3) and (4) with all $w_{i,j}$ set to unity.

Taking into account that the fMRI signal is sampled discontinuously according to the scanner repetition time (TR) the equations have to be rewritten in terms of a discrete-time model (see Dale, 1999). Based on an HDR of limited duration (T_{HDR}), the model for the fMRI signal is given as

$$\boldsymbol{y} = \boldsymbol{X_1 h_1} + \boldsymbol{X_2 h_2} + \ldots + \boldsymbol{X_{N_c} h_{N_c}} + \boldsymbol{n} \qquad (5)$$

where $\boldsymbol{y}$ is the vector of *Ntp* fMRI samples and $\boldsymbol{X_i}$ is the stimulus convolution matrix (SCM) representing the convolution of the event sequence with the individual HDR vectors $\boldsymbol{h_i}$ (dimension *Nh*, sampling interval for $\boldsymbol{h_i}$ assumed to be TR) and the noise vector $\boldsymbol{n}$ (dimension *Ntp*). Note that in contrast to Burock and Dale (in press) the event sequence is no longer a binary stimulus sequence but instead reflects the real-valued WF (see Equations (3) and (4)), which are thus embedded in the SCM. In case of short ISIs more than one event may occur during one TR period. The WFs of all corresponding events are then added and attributed to the corresponding position in the SCM. The initial values for all WFs are set to one, thus leading to the ordinary maximum likelihood estimate (MLE).

Assuming HDR models with a nonzero mean value the average overall $\boldsymbol{y}$ values observed over a run would also be expected to yield a nonzero average as well. The preceding trend correction would have subtracted this average. Therefore, in order to allow for HDRs with a nonzero average a global offset vector $\boldsymbol{o}$ is added resulting in the following equation.

$$\boldsymbol{y} = \boldsymbol{X_1 h_1} + \boldsymbol{X_2 h_2} + \ldots + \boldsymbol{X_{N_c} h_{N_c}} + \boldsymbol{o} + \boldsymbol{n} \qquad (6)$$

Concatenating all the individual SCMs plus the column of ones resulting from $\boldsymbol{o}$ to just one design matrix $\boldsymbol{X}$ and also concatenating the $\boldsymbol{h_i}$ vectors plus the offset value to just one $\boldsymbol{h}$ this equation may be reduced to (see Dale, 1999)

$$\boldsymbol{y} = \boldsymbol{X h} + \boldsymbol{n} \qquad (7)$$

$\boldsymbol{X}$ is of dimension *Ntp* by $N_c h$ where $N_c h = N_c Nh + 1$ specifies the length of $\boldsymbol{h}$, i.e., the total number of HDR coefficients to be estimated. In addition, the *Nw*

weighting coefficients have to be derived from the data, with Nw defined as

$$Nw = \sum_{i=1}^{N_c} Nw_i \qquad (7a)$$

Estimation of HDR Curves

Starting with all WF set to one, an unbiased MLE for $\boldsymbol{h}$ is provided by (see Dale, 1999; Hamilton, 1994)

$$\hat{\boldsymbol{h}} = (\boldsymbol{X}^T \boldsymbol{C}_n^{-1} \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{C}_n^{-1} \boldsymbol{y} \qquad (8)$$

assuming a zero mean Gaussian noise with covariance matrix $\boldsymbol{\Sigma_n} = \sigma_n{}^2 \boldsymbol{C_n}$ that may vary over voxels. In the following, however, we assume a simplified form ($\boldsymbol{\Sigma_n} = \sigma_n{}^2 \boldsymbol{C}$, i.e., a spatially varying variance with a fixed covariance structure) based on the observation of Burock and Dale (in press) that the overall performance of the estimation is not severely degraded using a global $\boldsymbol{C}$ matrix as compared to a local one. $\boldsymbol{C}$ is not known a priori and has thus to be estimated from the data, as outlined below.

The estimation covariance matrix is defined as (see Hamilton, 1994)

$$\boldsymbol{\Sigma_{\hat{h}}} = \boldsymbol{E}(\boldsymbol{e_{\hat{h}}} \cdot \boldsymbol{e_{\hat{h}}^T}) = \boldsymbol{s_n^2} \cdot (\boldsymbol{X^T C^{-1} X})^{-1} \qquad (9)$$

where the estimation error is defined as

$$\boldsymbol{e_{\hat{h}}} = \hat{\boldsymbol{h}} - \boldsymbol{h} \qquad (9a)$$

The statistical evaluation by means of an F test was described by Burock and Dale (in press), Dale (1999), Friston et al. (1998) (see also Hamilton, 1994).

IMPROVED ESTIMATION OF NOISE COVARIANCE

According to Burock and Dale (in press) the covariance $\boldsymbol{C}$ may be estimated by first deriving an intermediate estimate (according to Equation (8)) for $\boldsymbol{h}$ assuming a trivial covariance (i.e., $\boldsymbol{C} = \boldsymbol{I}$, the unit matrix), then estimating the empirical autocorrelation from the residual errors averaged over all voxels of the brain volume and finally fit the parameters of the following model (Zarahn, Aguirre, & D'Esposito, 1997; Weisskoff et al., 1993):

$$K(n) = \begin{cases} 1 & n = 0 \\ (1 - \lambda)\rho^n & 0 < |n| \leq R \\ 0 & |n| > R \end{cases} \qquad (10)$$

With TR = 2 sec we set R to 20 as a trade-off between computational effort and accuracy.

It should be stressed that the covariance structure of the residuals differs from the underlying noise process. This is because the design matrix induces model specific correlations in the residuals that are not present in the original noise process. To test for potential effects on statistical inference, Burock and Dale (in press) used this covariance model to estimate the structure of synthetic as well as actual fMRI noise data. They observed that the false-positive rates of the HDR F test were appropriately controlled (as opposed to the results derived under the assumption of white noise). These results indicate that the model is a reasonable approximation leading to a significant improvement of the HDR estimation.

In contrast to Burock and Dale (in press), in our implementation approach λ and ρ are fit directly with a nonlinear steepest gradient optimization procedure rather then taking the logarithm of $K(n)$ and estimating the two parameters running an ordinary least square fit. This modification leads to a better estimate because it is not restricted to positive correlation values and at the same time leads to greater robustness with respect to noise.

$\boldsymbol{C}$ is derived from the modeled autocorrelation function by building a Toeplitz-type matrix from shifted versions of the autocorrelation.

ASSESSMENT OF AMPLITUDE VARIATIONS — JOINT NONLINEAR OPTIMIZATION OF WF AND HDR

In order to allow for variable HDR amplitudes, the design matrix is no longer restricted to zeros and integers (indicating the occurrence of stimuli), but instead the design matrix may be modified with the goal of optimizing the linear model with respect to the WF as described in Equations (3), (4) and (4a). From the combination of Equations (3) and (4) it follows immediately that a combined estimation of the HDRs and the WF leads to a nonlinear equation that can no longer be solved directly by a MLE approach. An iterative least squares procedure may be applied to solve this problem. This involves using the linear estimator of the HDR given the current estimates of the WFs and then using the same estimator to estimate the WFs given the current estimates of the HDR. This procedure has shown to converge rapidly in case of simple models (for instance one HDR, three WFs) with low temporal overlap of subsequent HDRs. However, with more complex models and severe overlap as discussed here, the performance is poor and another algorithm is needed to speed up the convergence. Therefore, we developed a nonlinear steepest descent iterative optimization scheme as outlined in the following.

Starting from the initial $\boldsymbol{X}$ matrix and the corresponding MLE estimate as an initial guess for $\boldsymbol{h}$ the WF and $\boldsymbol{h}$ are optimized by applying a modified steepest descent algorithm as proposed by Powell (see Press, Flannery, Teukolsky, & Vetterling, 1988). The optimization is

guided by the criterion of minimizing the residual error s_y^2, which is

$$s_y^2 = \| \boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{b}} \|^2 \quad (11)$$

The following steps are repeated recursively until the error criterion stops the iteration:

(A) Determine the direction of steepest descent at intermediate solution $\boldsymbol{WF_i}$ (i.e., the vector carrying the set of WFs) based on $\boldsymbol{h_i}$. At the first loop, the WFs are all set to one and $\boldsymbol{h_i}$ is set to $\boldsymbol{h}$ as discussed before.

For this purpose, each of the WFs is modified independently by a certain step. With each modification the resulting MLE estimate for vector $\boldsymbol{h_i}$ is derived and from that the current residual error s_y^2.

(B) $\boldsymbol{WF_i}$ is modified along the steepest descent direction by subsequent steps until the three latest $\boldsymbol{WF_i}$ vectors cover the interval where the intermediate minimum residual error is located. Usually this condition is met after two to four steps. With each step, the step size is doubled.

(C) In order to find the $\boldsymbol{WF_i}$ with smallest residual error along the current gradient, a second-order polynomial is fit to the residual errors obtained during the last three steps. The $\boldsymbol{WF_i}$ corresponding to the minimum error is derived as the coordinate vector of the minimum of the polynomial.

(D) For each of the intermediate $\boldsymbol{WF_i}$, the corresponding $\boldsymbol{h_i}$ is estimated applying the MLE algorithm. The residual error mentioned in (A) and (C) is evaluated based on this updated $\boldsymbol{h_i}$ vector.

(E) Keep this intermediate optimum vector $\boldsymbol{WF_i}$, reduce step size by a factor of 2, and go back to (A).

The sum of the WFs per event type is always set to Nw_i (number of WF/event type) in order to reduce the degrees of freedom and at the same time to prevent the iteration from running into instability. In addition, the values for the WFs are restricted to the interval [0,2]. The optimization is stopped if either the number of iterations has reached a preset maximum or if the decrement between subsequent iterations or the variation of the parameters stays below another predefined limits over two subsequent loops. In order to avoid local extrema the solution subspace is scanned with respect to the residual error by a Monte Carlo method once the algorithm has found a minimum. At the same time, the determinant of the Hessian matrix of second partial derivatives provides a test for uniqueness of the solution.

In the subsequent text, the extended approach will be called weighted MLE (WMLE).

Error of HDR and WF estimate

Given a normal distribution for the non-linear estimates (see Experiments section for a corresponding test) the covariance of the non-linear HDR estimates can be derived by Equation (9) (covariance for the linear HDR estimate). Conversely, the MLE formalism as specified above can be applied for a maximum likelihood estimation of for $\boldsymbol{WF_i}$ given the actual HDR estimates. This means that the $\boldsymbol{h_i}$ vector defining the HDR is included as a set of constants in the SCM and $\boldsymbol{WF_i}$ finally estimated according to Equation (8). With this inverse MLE scheme, Equation (9) now specifies the approximated error covariance for the WF estimates.

SIMULATIONS

The main goal of the simulations was to test with well-defined data (i.e., known time course of activation and known noise structure) if (i) the WMLE procedure would be able to deconvolve different HDRs even in case of severe overlap and rapid stimulation paradigms and (ii) to prove that WFs can be reliably extracted. In two of the three simulation studies, we also derived the MLE estimates based on separate HDRs (instead of one HDR plus separate WFs) for comparison.

TR was always assumed to be 2 sec. The number *Nh* of samples per HDR was set to 9 corresponding to a 16-sec duration (starting at stimulus onset) based on a sampling interval of 2 sec for the HDR. The possibility of defining the HDR with sampling intervals other than the TR will not be addressed here.

Simulation I: One HDR With Three Different WFs

This simulation served to prove that the nonlinear optimization would extract correct WFs from a data set with HDRs exhibiting varying amplitudes. In addition, the improvement WMLE provides in terms of estimation error was to be compared to the MLE scheme.

Materials and Methods

A synthesized fMRI data set was generated assuming a balanced randomized sequence of 1182 HDR, which were identical in shape but modulated by three different WFs (0.6, 0.9, 1.5). The resulting data sequence was overlayed on a synthetical noise data set generated by a colored noise process (temporal correlation according to Equation 10) with model parameters $\lambda = 0.75$ and $\rho = 0.88$ as specified by Burock and Dale (in press). The ISIs were jittered over an interval of 800 to 1200 msec in order to improve the estimation efficiency as outlined by Dale (1999) The HDR shape was defined to resemble usually observed empirical response functions (Dale & Buckner, 1997; Boynton et al., 1996). Simulations were repeated for 300 data sets of 1300 samples each with independently generated noise. Three different settings were applied according to the signal-to-noise ratio (SNR) set to 2.0, 0.9 and 0.2 respectively. SNR was defined as

$$SNR = \| \boldsymbol{Xb} \|^2 \; / \; \| \boldsymbol{y} - \boldsymbol{Xb} \|^2$$

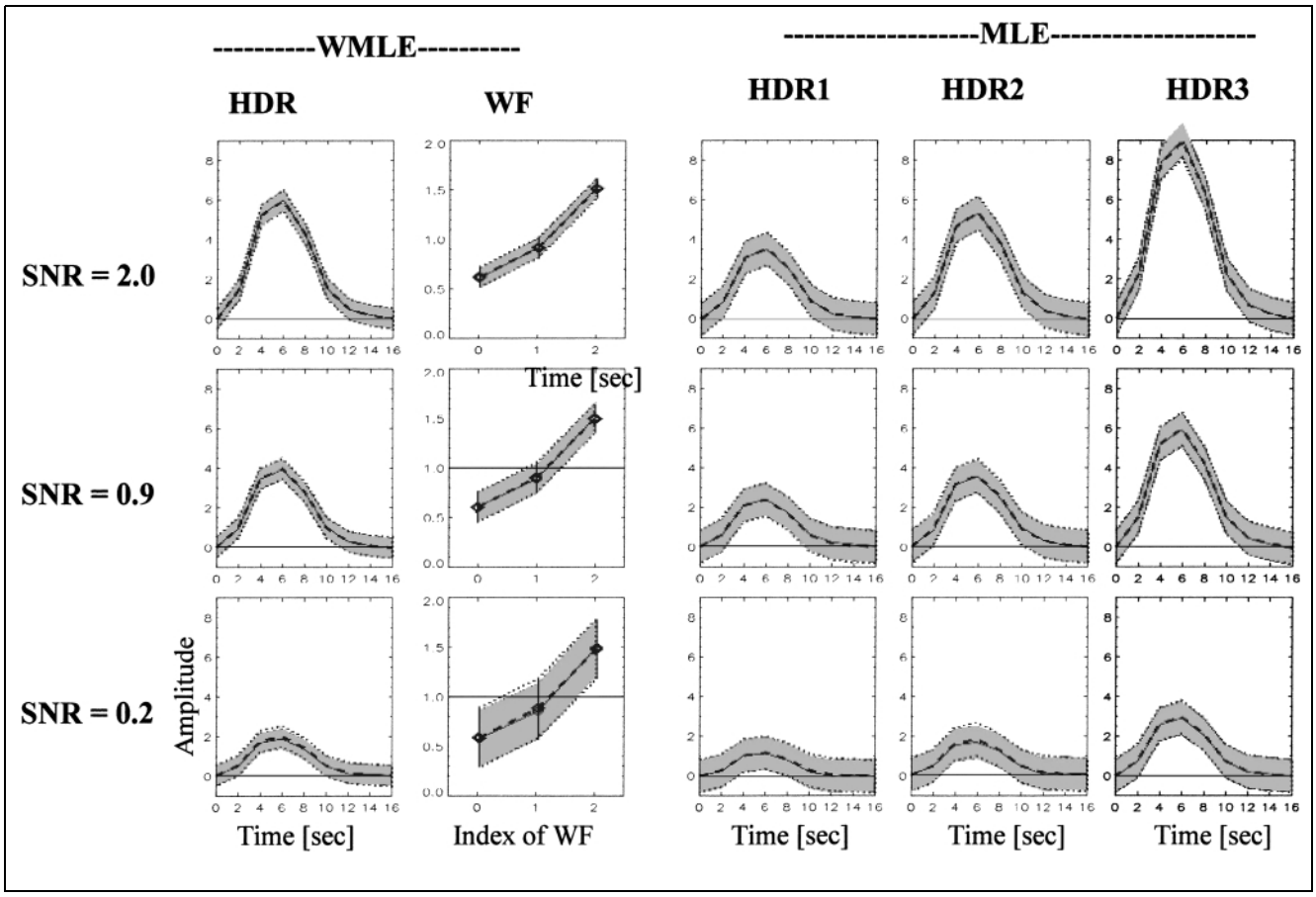

Figure 2. Results of HDR estimation of synthetic data with HDR of fixed shape and three different amplitudes. Conventional MLE versus WMLE method at different SNRs. Results derived from 300 repetitions (separate data sets). Solid line: average estimated HDR and WF (marked by rhombs). Dashed line: true HDR and WF used for signal generation. Shaded region: empirical standard deviation. Dotted lines: theoretical estimation error as derived from Equation (9), centered around true HDR.

Results

The WMLE estimates for the HDRs as well as the three estimated WFs are plotted in Figure 2 together with the original time course assumed for the HDR during signal generation. For all three SNR values the estimated HDR resembles the original one within a small range representing the empirical estimation error, which in turn closely resembles the estimation error according to Equation (9) (see Figure 2). In addition, the MLE estimates are also plotted based on a model with three different HDRs. Comparing the two results, the reduced HDR error range of the WMLE model is clear. The average standard deviations of the two MLE and WMLE estimates are listed in Table 1.

Simulation II: Test of Statistical Efficiency

In order to evaluate the gain of WMLE over MLE in terms of statistical efficiency we evaluated the statistical detection rates in a data set with artificial activations overlayed on noise.

Materials and Methods

The pool of synthesized data sets from simulation I was modified to SNRs of 0.057, 0.128, 0.228 and 0.513. In addition, another 3000 time courses were generated containing colored noise only (of same variance), i.e., without any superimposed response function. In order to evaluate the statistical efficiency we applied both estimation schemes (one HDR plus three WFs using WMLE as opposed to three separate HDRs using MLE) to the whole data set. Applying the *F* test to all 300 and 3000 estimates according to Burock and Dale (in press)

Table 1. Empirical Estimation Error Variance Derived from Estimations in Simulation I for Different Signal-to-Noise Ratios. Average Overall Samples of Estimated HDR and WF, Respectively

	SNR = 2.0	*SNR = 0.9*	*SNR = 0.2*
MLE: HDR	0.696	0.693	0.695
WMLE: HDR	0.273	0.272	0.256
WMLE: WF	0.009	0.022	0.082

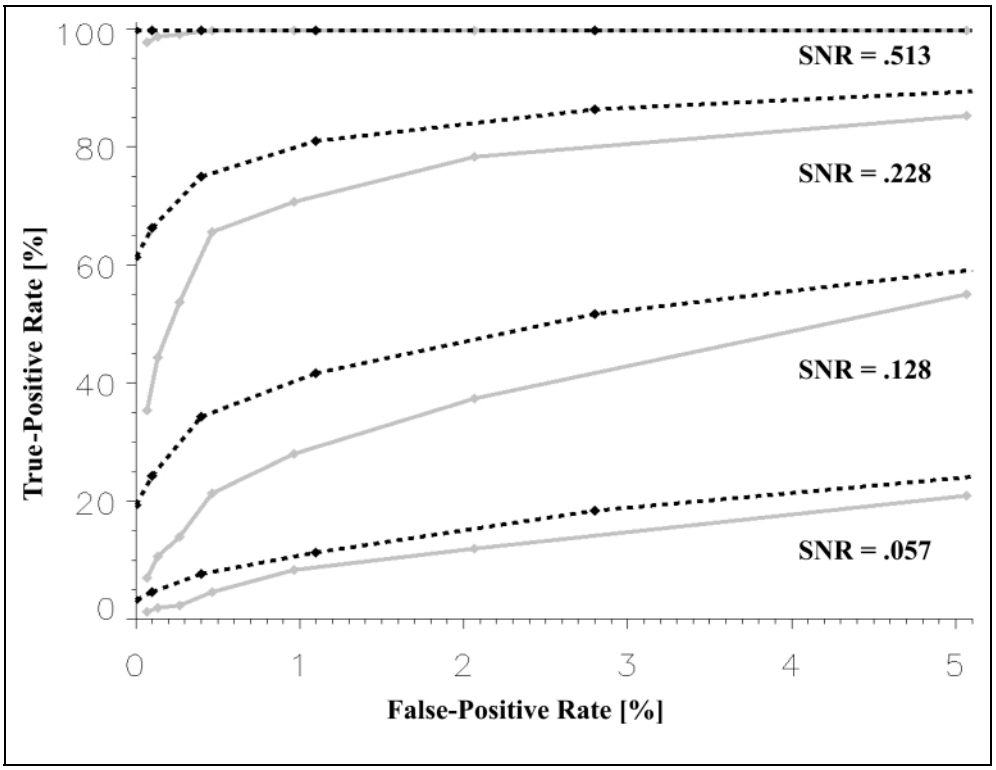

Figure 3. ROC obtained using MLE (solid line) and WMLE (dotted line) method to estimate HDRs. True and false positive rates of activations detected by an *F* test are plotted for different signal to noise ratios (SNR) derived from artificial colored noise overlayed by artificial activation.

we determined the true-positive and false-positive rate, respectively (assuming the null hypothesis), and thus estimated the receiver operating characteristic (ROC; see Xiong, Gao, Lancaster, & Fox, 1996). These tests were performed for all four SNRs.

Results

The results in terms of ROC curves are plotted in Figure 3 for both estimation schemes calculated for significance levels α in the range of $0.0001 < \alpha < 0.02$. The WMLE procedure always yielded better performance in terms of statistical efficiency. The relative enhancement increased with SNR. From the ROC curve, it can be derived that the gain in statistical efficiency was especially strong at small values of α.

Simulation III: Combined Estimation of Short Term HDR Overlapped by a Sustained HDR

In experiment II (see below) we tested if the estimations of both the HDR and the correct WFs would empirically yield reliable results even in case of severely overlapping HDR of different shape. The present data set aimed at simulating such an experiment and evaluating the performance of the WMLE approach with known HDRs and WFs.

Materials and Methods

Another 100 data sets were generated according to the procedure applied for simulation I, however, with two different HDR shapes. One of them (''HDR1'') represents a response with a very early onset and a decay at about 6 sec whereas the second ('HDR2') was characterized by a response onset at about 3 sec but an extended duration of about 16 sec, thus reflecting a type of sustained activation. This experiment aimed at exploring whether both the MLE and the WMLE algorithm would be able to extract both a sustained and a phasic activation even in an experiment in which a serious temporal overlap of the two HDRs might occur. The different HDRs were convolved with a stimulation sequence consisting of 216 frames each comprising an initial HDR2 stimulus followed by 10 HDR1 stimuli with ISIs jittered over a range of 800–4000 msec (approximately Poisson-distributed, mean 1000 msec). The interval between HDR2 and the first HDR1 was 4 sec in 80% of the frames and 1 sec in 20%. In order to check the ability of the algorithm to correctly extract WFs under these conditions we attributed two different WFs (0.7 and 1.3) equally distributed to the subsequent HDR2 in a balanced order. The study was run for three different SNRs (1.3, 3.6, 7.1).

Results

Both the usual short-term HDR as well as the extended version were correctly extracted from the simulated data set within an error range predicted by Equation (9). In addition, the WF derived with the WMLE met the theoretical values, see Figure 4. This means that even in case of an extreme temporal overlap the two HDR characteristics can be distinguished well and amplitude

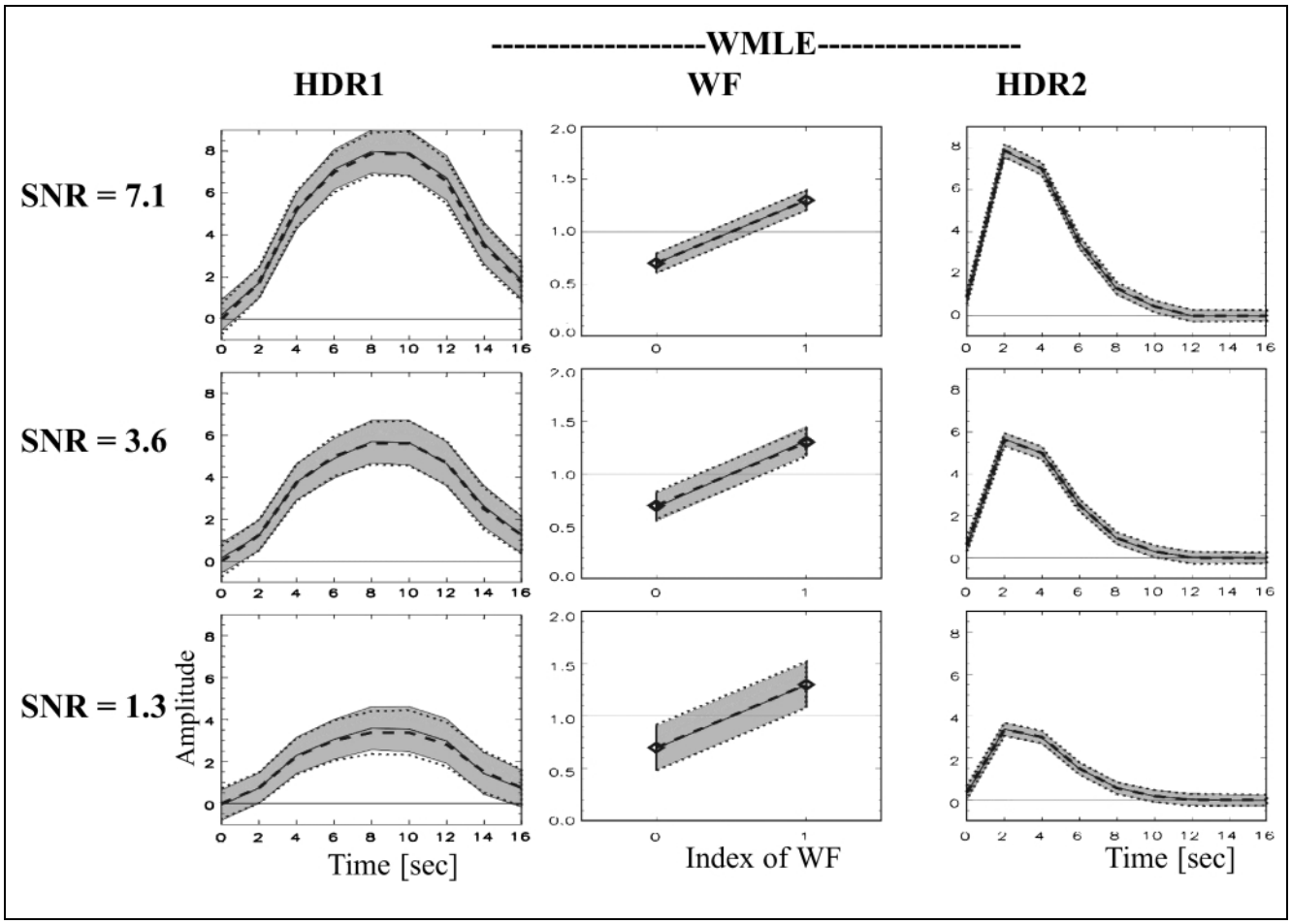

Figure 4. Estimation result derived with WMLE method from 100 artificial data sets comprising a superposition of two HDR with different temporal shapes. HDR1 was included with two different WFs as indicated by the second column. Results derived from 100 repetitions (separate data sets). Solid line: average estimated HDR and WF (marked by rhombs). Shaded region: empirical standard deviation. Dashed line: true HDR and WF used for signal generation. Dotted lines: theoretical estimation error as derived from Equation (9), centered around true HDR.

modulations can be determined given appropriately randomized ISIs.

Discussion of Simulations

In simulation I it was shown that even in case of severely overlapping response functions (ISI about 1 sec, HDR duration about 10 sec) the new WMLE procedure is able to extract both the HDRs and also the WFs reliably within the error range predicted by the corresponding MLE theory. It is worth mentioning that similar results were observed with a variety of synthetic HDR shapes tested. As predicted from the MLE formalism the error range increases with decreasing SNR, however, without leading to unreasonable results thus empirically proving the stability of the nonlinear optimization. As outlined by Dale (1999) (see also Burock, Buckner, Woldorff, Rosen, & Dale, 1998; Woldorff, 1993) jittering the ISIs is a prerequisite for an efficient deconvolution procedure. This fact is of paramount importance if subsequent HDRs of different shapes are severely overlapping as specified in simulation III. As can be seen from Figure 4 the sustained HDR starting at the beginning of each frame overlaps with almost all of the subsequent HDRs of the frames. Without jittering the ISI between all subsequent events, neither the different HDRs nor the corresponding WFs would have been extractable.

The development of the WMLE algorithm was initiated by the goal of effectively monitoring fMRI activations in experiments where a constant-shape HDR with fluctuating amplitudes can be assumed. As a side effect, the statistical efficiency is improved due to the reduced number of variables to be estimated. This theoretical consideration was proven in simulation II. Experiments where only slight activations are expected, therefore, might benefit from the improved statistical sensitivity of the WMLE algorithm provided the experiments meet the requirements needed to apply the WMLE algorithm.

The artificial data used in these simulations had a noise distribution that was modeled to closely match the noise structure observed with real fMRI data (see Burock & Dale, in press; Zarahn et al., 1997). The WMLE algorithm applying this model may lead to overoptimistic conclusions, because with real data this model may be less

appropriate and thus cause less stable results. The following fMRI experiments aimed at further validation.

EXPERIMENTS

In this section we present two fMRI experiments demonstrating the additional information and the improved performance provided by the WMLE approach in situations when HDR of constant shape but variable amplitudes may be assumed.

Experiment 1: Visual ErfMRI at Three Different Visual Contrast Ratios

This experiment was included in order to check the estimated WF by some physiological plausibility. As known from ERP experiments (see for instance Petry, Donovan, Moore, Dixon, & Riggs, 1982) and also from fMRI studies (see Goodyear & Menon, 1998) a monotonic relation between cortical activation and physical properties would be expected in this kind of experiment. Accordingly, the WF should reflect this relationship. In addition, the modulation of the WF should mimic the main amplitude variations of the MLE estimate based on three different HDRs.

A further goal of this experiment was to test empirically if despite the nonlinear optimization the distribution of the WMLE–HDR estimates would match the normal distribution assumed for the MLE estimates.

Materials and Methods

One subject underwent a visual perception fMRI experiment applying stimuli of identical shapes but of three different contrast ratios. Stimuli consisted of a black and white spatially smoothed (Gabor profile) checkerboard (size of stimulus 18.6° × 14.0°, grid spacing 1.6°) that was globally modulated according to a sine wave profile (largest contrast in central stimulus decreasing to zero at the borders). These were presented for 100 msec at the center of the field of view using a visual presentation system in which the images were projected onto a back projection screen in front of the subject. A fixation cross was superimposed on the center of the stimulus. The subjects' task was to monitor and count infrequent changes of the fixation cross ('+' turning into '×'). The stimuli were presented at three different contrast ratios (low = 36%, medium = 78%, high = 100% at the center of the stimulus). Also, a null event ('no stim') according to the definition of Burock et al. (1998) was added in order to be able to compare the MLE and the WMLE estimates of the HDRs to those derived via selective averaging and subsequent subtraction (these results not reported here). The first-order sequence of these four event types was fully counterbalanced. A total of 1182 stimuli were presented with the ISI jittered over the range 1000 to 2000 msec.

fMRI signals were acquired during visual stimulation from one subject with a General Electric 1.5-T Signa LX neuro-optimized scanner running an echo-planar gradient echo sequence (TR = 2 sec, TE = 40 msec, flip angle 90°). A standard head coil was used to pick up the rf signal. Slice thickness was 7 mm (1 mm gap) with an in-slice resolution of 3.125 × 3.125 mm (64 × 64 matrix). The slices were oriented perpendicular to the calcarine fissure. A total of 1300 (split into 10 runs of equal length) images were acquired. Before further analysis the data were slice–time corrected and spatially filtered applying a 3-D Gaussian filter with a 6 mm full width half maximum (FWHM).

As specified in simulation I we applied two different approaches to estimate the HDR: (i) Individual HDRs for each of the three different stimulus contrasts (i.e., the regular MLE estimate) and (ii) one common HDR but three different WFs (i.e., the WMLE estimate). These estimates were derived from voxels located within the center of activation as identified by the F test (with Bonferroni correction) applied to the MLE estimates.

With respect to statistical inference the normality assumption of the WMLE–HDR estimates for fMRI noise data was evaluated as follows: 3000 contiguous voxels were selected from a 3-D rectangular area covering the central sulcus (which is located clearly outside the activated regions according to the MLE procedure). The raw data observed in these voxels were analyzed using both the MLE and WMLE algorithm, assuming just one HDR for both procedures and three WFs for the WMLE–HDR. The event sequence applied here was not synchronized with the original one. The empirical distribution of the MLE and WMLE estimates were normalized with respect to the standard deviation and then subjected to a two sample Kolmogorov–Smirnov test (KST).

Results

The statistical evaluation of the MLE estimate yielded an activation cluster with the center located in V1 close to the midline but also extending to more lateral structures, see Figure 5. From both regions representative voxels were randomly selected for further analysis with the WMLE procedure. The resulting WMLE–HDR and WF estimates are plotted in Figure 5 together with the three MLE–HDR estimates. Again, a comparison of the two estimation techniques clearly shows the reduced estimation error for the WMLE approach. In addition, the WFs increase monotonically with the contrast. The same tendency can also be derived from the three separate MLE estimates, but at the cost of a larger estimation variance. As marked in Figure 5, the ratio of peak amplitudes (defined as the maximum of a second-order polynomial fitted to the three largest HDR samples) of the MLE– and WMLE–HDRs are close to the WF and thus illustrate the relation between MLE and WMLE.

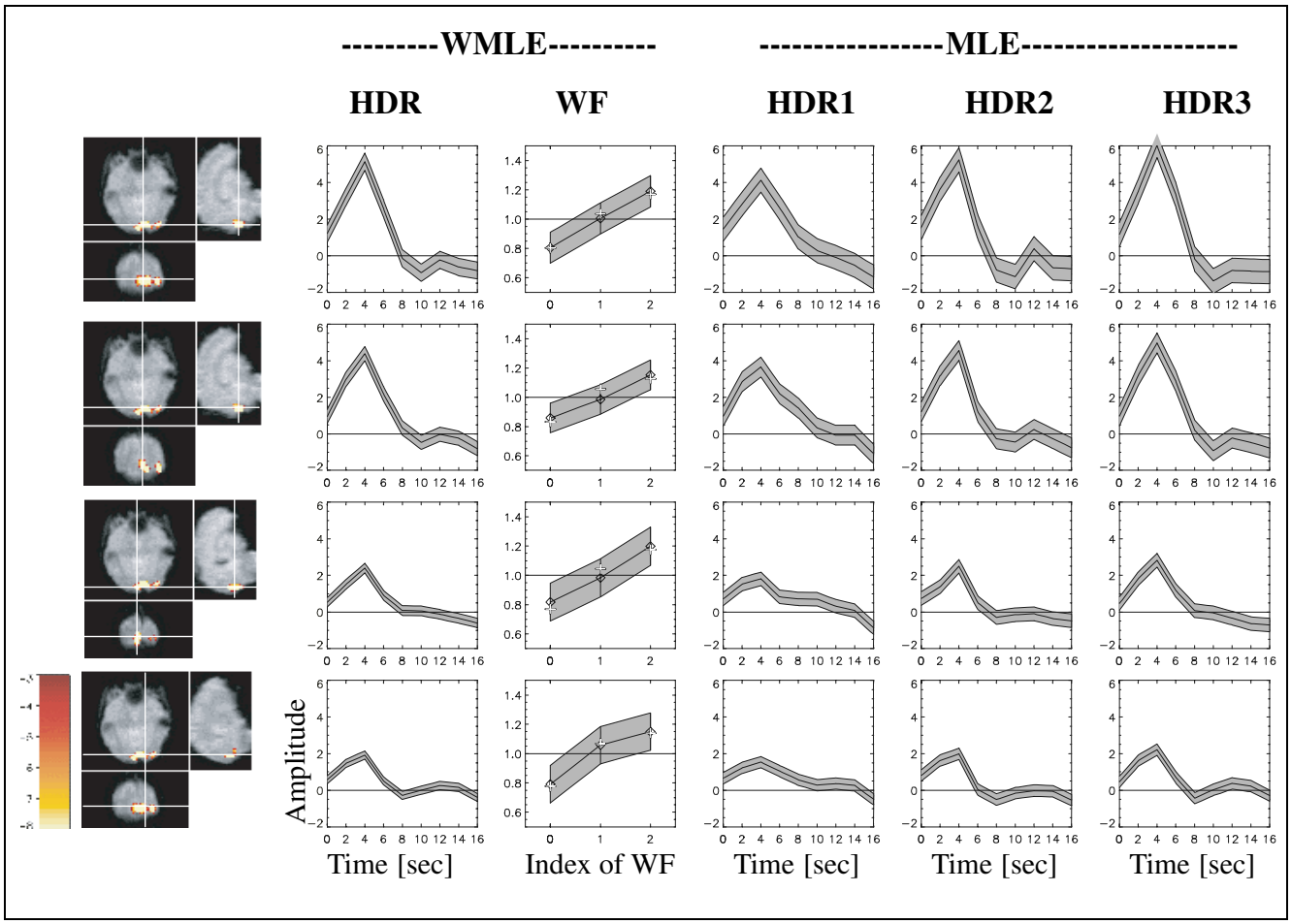

Figure 5. WMLE estimate of an fMRI experiment with visual stimulation with Gabor pattern presented at three different contrast ratios. Results shown for four voxels. First column: significant activation ($P < .001$ after Bonferroni correction) displayed according to transaxial (upper left), sagittal (upper right), and coronal (lower left) slice. The locations of the four voxels are indicated corresponding to the results displayed in rows 1–4. The indices of WF correspond to contrast ratios of the stimuli (0, 1, 2 = 36%, 78%, 100%). Solid line: estimated HDR and WF (also marked by rhombs). The crosses overlayed on the WF rhombs indicate the MLE-HDR amplitudes normalized with respect to WMLE-HDR peak amplitude. Shaded region: estimation error as derived from Equation (9), centered around estimated HDR.

Regarding the distributional properties of the WMLE estimates, the KST did not reject the hypothesis of equal distributions at a level of $P < .05$ at all samples of the estimated HDR.

Discussion of Experiment I

Keeping in mind the earlier reports on the relation between sensory stimulus variations and neural activity, the assumption of a fixed HDR shape with varying WFs was justified. In this sense the WMLE approach is empirically validated according to the following observations (see Figure 5): (i) The WFs closely resemble the overall amplitude variations of the three HDR estimated by the classical MLE approach, and (ii) the WF monotonically increase with the contrast ratio. Also, the result that similar WF gradients were observed in different voxels (even in those separated by more than the FWHM of the 3-D filter) of the activated area provides further evidence that the analysis reflects physiological modulations rather than chance effects.

From the result of the KST it can be concluded that the normalized distribution of the WMLE estimates is identical to the one observed for the MLE estimates (which is supposed to be normal). This means that the nonlinear procedure does not significantly affect the distribution of the estimated HDR. We conclude, that WF estimates are also normally distributed, taking into account that in the last step of the WMLE algorithm the WFs are derived by application of the linear MLE.

Experiment II: Sequential Effects in Visual Spatial Attention

Event-related brain potentials (ERPs) have revealed various sequence effects in the neural responses to attended or unattended stimuli, e.g., the amplitude of the early sensory components (P1, N1) elicited by a stimulus at the attended location varies depending on

the direction of attention with respect to the preceding stimulus (Woldorff & Hillyard, 1991; Luck, Heinze, Mangun, & Hillyard, 1990). An important question, then, is whether similar sequence effects can be observed for the HDR in a spatial attention paradigm. In the following fMRI study, a spatial cue was followed by a sequence of bilateral visual stimulus arrays, and the subject had to discriminate a target item from surrounding distractor items at the attended location. The deconvolution approach described above was used (i) to assess possible shape differences of the overlapping HDRs related to the cue and items, and (ii) to detect amplitude variations of the HDRs with respect to the sequence of the direction of attention as indicated by the cue, i.e., LL (attention left followed by attention left) or RR versus LR or RL.

Materials and Methods

Six subjects were included in the study. Images were acquired using a General Electric 1.5-T Signal LX neuro-optimized scanner. After structural scanning (IR-prepared EPI) functional images were acquired in nine runs lasting 8.5 min each (TR: 2 sec, TE: 40 msec, flip angle: 90°, matrix: 64 × 64 voxels, 2.8125 × 2.8125 × 3 mm, 14 slices oriented vertical to the calcarine fissure). Rf signals were picked up with a 5-in. surface coil placed under the occipital part of the head. A high-resolution structural scan was acquired in a separate session applying a 3-D SPGR sequence (124 slices/1.5 mm thickness, 256 × 256 matrix/FOV 25 cm). Using SPM 99 (Wellcome Department of Cognitive Neurology, UK) the raw fMRI data were movement and slice-time corrected, matched into Talairach space (Talairach & Tournoux, 1988), resampled to a 2 × 2 × 2 mm matrix, and finally spatially filtered with a 3-D Gaussian filter (6 mm FWHM).

A symbolic cue indicating right, left, or neutral was presented for 500 msec followed by a pause of 500 (20%) or 3500 msec (80% of trials). Subjects were asked to covertly shift their attention to the side indicated by a cue. Then, a sequence of 10 bilateral stimuli ('item') appeared in the upper visual field (eccentricity 8°, 1° above horizontal meridian, 1.2 × 1.2 degrees extent, duration 200 msec, ISI 800–4000 msec, mean 1000 msec, Poisson-distributed). Stimuli on both sides consisted of 3 × 3 arrays of crosses superimposed on a background checkerboard that was locally and globally smoothed. The central element in the array was a T that could be displayed upright or inverted (50/50%).

The subjects had to discriminate whether the T in the middle of the attended stimulus was upright or inverted and to press one of two buttons. During the neutral condition, they were told not to shift attention but to detect the appearance of each bilateral stimulus without discrimination. Potential eye movements were monitored by means of a home-made control system with a fiber optic device using infrared light (resolution approximately 1°). During each run 24 blocks of the three conditions (plus six blocks of the fixation condition) were presented in a counterbalanced order lasting 18 sec each.

Using the MLE algorithm six different HDRs were attributed to the event types CL, CR, CN, IL, IR, IN, where C = cue, I = item, L = attention directed to the left hemifield, R = attention directed to the right hemifield, and N = neutral condition). Attention-specific clusters of significant voxels were determined running an *F* test (Bonferroni corrected) (see Burock & Dale, in press) for the left minus right difference cue HDR. Significant activations were found in the fusiform gyrus, as originally reported by Heinze et al. (1994) in this sort of lateralised attention task for upper-field bilateral stimuli. We focused on this area in order to trace attention-related sequence effects (as discussed above). For this purpose, a few representative voxels located in the activated part of the right fusiform gyrus were selected. From the corresponding data sets, HDRs and WFs according to the WMLE procedure were estimated; however, in this case only two different HDR shapes (cue and item) were assumed. For both HDRs, individual sets of WFs were derived for the responses occurring at the LL, RL, NL, RR, LR, NR, and *x*N (*x* representing an arbitrary predecessor) sequence of attention conditions. Regarding the 10 items following each cue a common WF was used, i.e., no temporal variations within a run were accounted for here.

Results

In the following we show the results observed in a representative voxel from the right hemisphere of a representative subject.

In the left part of Figure 6 the HDRs and their corresponding standard deviation ranges as estimated with WMLE are shown (i) for the cue and (ii) for the item. On the right side, the WF as derived for the different sequences are displayed (including confidence regions).

WFs were generally smaller for cues and items at ipsilateral than at contralateral condition. In the attention left condition (i.e., contralateral) the WF according to a RL sequence was much larger as compared to LL. Applying a *t* test to the WF we looked for intraindividual statistically significant differences between the same–same cue WF and the same–different cue WF. For the voxel shown here the difference was significant at a level of $p < .01$. Throughout all voxels of the significant cluster this difference was significant at least at a level of $p < .05$. In addition, the NL sequence showed a difference compared to LL (not shown in the figure) almost as large as RL. The corresponding item related WFs exhibited a similar tendency in some voxels without reaching any level of significance. Only minor differences were encountered for other sequence-related activity,

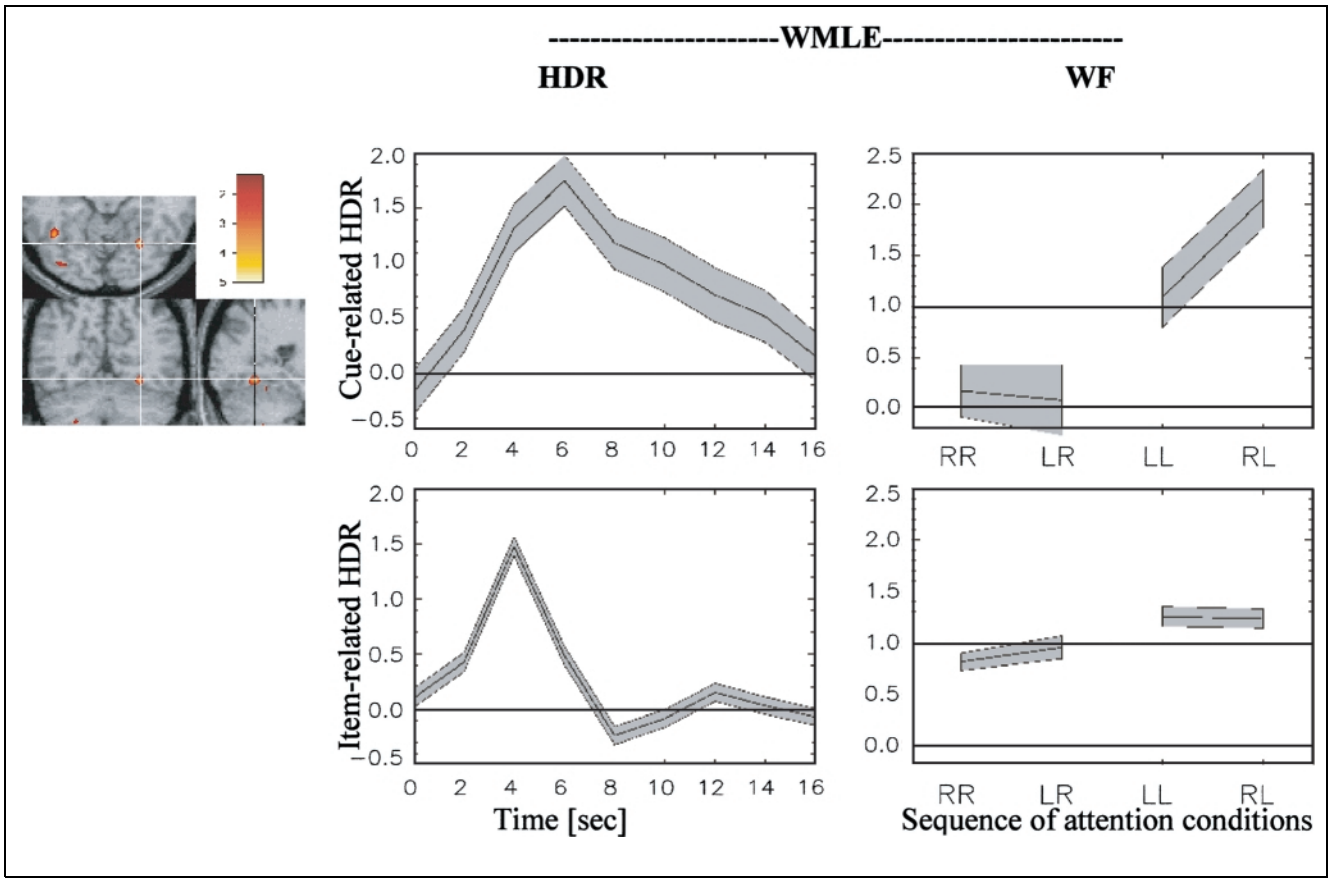

Figure 6. WMLE estimate of an HDR and WF for a visual spatial attention fMRI experiment. The HDRs for the cue and the item differ due to the different numbers of occurrences. Upper left: cluster of voxel ($P < .05$ after Bonferroni correction) exhibiting significant difference specifically on attention direction. First column: separate item- and cue-related HDR estimated by WMLE. Solid line: estimated HDR. Shaded region: estimation error as derived from Equation (9), centered around true HDR. Second column: separate WFs derived for same–same and same–different sequence of attention direction. Solid line: connecting two estimated WFs, i.e., indicating difference between different sequences. Shaded region: estimation error as derived from Equation (9), centered around estimated WF.

none of these showing a stable pattern over all voxels included.

Discussion of Experiment II

Luck et al. (1990) reported that ERP components observed during a visual spatial attention task decreased if two subsequent targets occurred in the same hemifield. Based on behavioral measures Ward, Duncan, and Shapiro (1996) put this effect in a more general concept and showed that substantial interference may occur between sequence and mental processing of subsequent visual objects. Accordingly, we made the assumption of (i) separate HDR shapes for cue and item and (ii) separate WFs for different sequences of attention directions.

As shown in Figure 6 the WMLE algorithm extracted different HDRs for cues and items. Also, under the contralateral condition (i.e., attention to the left hemifield) a clear-cut modulation (t test, $p < .01$) of the sequence-related WF was observed for the cue-related HDR but not for the items. The significant difference between WF determined for a same–same sequence as compared to same–different resembles the abovementioned findings of Luck et al. (1990). No clear-cut results were found for the HDRs estimated for those cues drawing attention to the ipsilateral field of view. In addition, no significant effects were observed with the item related amplitude. This statement, however, only holds for the small data set reported here, in this paradigm: one common WF for all 10 items following the cue. It cannot be excluded that temporal fluctuations within the run are hidden by this way of analysis (see Hopfinger et al., 2000). A detailed analysis with respect to the temporal structure of the within-run items could uncover those effects.

GENERAL DISCUSSION

In the present paper we describe a special version of the general approach to extract HDR functions from raw fMRI data based on the assumption that the HDRs follow a linear superposition model. With respect to the more general approach of Dale (1999), which requires the

estimation of a large number of variables in case of multiple different HDRs, the goal of the present study was to find a modified version that is more parsimonious with respect to the number of these variables.

With appropriate simulations we have tested the new approach regarding its capability of reliably extracting different HDRs and associated WFs. In addition, the gain in statistical efficiency was demonstrated. Experimental data from two fMRI studies were analyzed in order to test the algorithm under controlled physiological conditions and secondly to provide a good example of the application in the context of cognitive experiments.

A general way to reduce the complexity is to impose constraints restricting the application to certain special cases. For instance, the approach of Josephs et al. (1997) and Friston et al. (1998) restricts the HDR to linear combinations of a small number (two or three) of basis functions thereby providing an efficient way of defining (and thus estimating) the HDR. However, this formalism is limited to data sets that can be adequately described by assuming this restricted range of HDR shapes. In order to overcome this limitation at least for some experiments we formulated a different, less restrictive constraint requiring the HDR to be constant in shape for a class of events (although not being restricted to a predefined class of forms) and allowing for nonstationarity of the amplitudes. This approach is principally suited to physiological experiments where sequence or context effects are likely to modulate the amplitudes of neural responses. In addition to the experimental examples given before in this paper, other experiments similar to the oddball paradigm may be candidates for WMLE.

In order to bring further down the false positive rate in statistical inference we accounted for temporally correlated noise by including the noise covariance in the estimation scheme (Zarahn et al., 1997) as proposed by Burock and Dale (in press). These authors have demonstrated that statistical efficiency significantly benefits from a proper estimation of the individual statistical noise structure. Therefore, we slightly modified their estimation procedure and made it more robust. In addition, our algorithm accounts for both positive and negative correlations.

As a consequence of the parsimonious number of variables, the estimation error is reduced and thus the statistical sensitivity is improved (see simulation study in this paper), provided the assumption of a stationary noise structure is true. However, taking into account the considerable computational effort added by the nonlinear optimization this advantage can not always be exploited for practical reasons. For instance, for experiment I it takes about 1 min/voxel on a SUN SPARCstation 5 computer to estimate the HDRs and the WFs with the WMLE algorithm. Therefore, for routine applications a reasonable trade-off might be first to define a cortical area by means of a conventional fMRI analysis and then focus the advanced analysis on this reduced number of voxels.

Comparing the different linear approaches addressed here, the specific fields of application are mainly defined by physiological and experimental considerations. If no evidence at all is available regarding the time course and variation of the HDR the general unbiased estimation approach of Dale (1999) and Burock and Dale (in press) will be most appropriate as it does not require any restriction regarding the HDR time course (except to be of finite duration) and thus avoids missing or mischaracterizing activations due to an oversimplification. If stronger assumptions regarding a certain shape and onset of the HDR is a priori reasonable, more efficient estimates may be achieved by expressing the HDR functions as a linear combination of a few basis functions, as described in Josephs et al. (1997), Friston et al. (1998), and Dale (1999).

Finally, if context or sequence effects are likely to modulate the amplitude of neural response, the new algorithm presented in this paper will be useful to improve the estimation and trace the temporal variation. In any case, the final decision regarding the choice of analysis has to be made along plausibility considerations because the shape and variability of the real HDR cannot be predicted in absolute terms.

CONCLUSION

We have described an extension to the framework of linear estimation of event-related HDR functions aimed at accounting for sequence effects by introducing WFs that modulate the estimated HDR. The new algorithm comes out with a reduced estimation error due to the smaller number of variables thus leading to an enhanced statistical efficiency. Analyzing both synthesized data sets and real fMRI data has demonstrated that (i) the algorithm yields reasonable results even with short ISI given an appropriate jittering, and (ii) sequences of HDRs with strongly differing duration and a correspondingly severe overlap can be disentangled with respect to varying HDRs and WFs.

The new algorithm is thus suitable for the evaluation of arbitrary experiments dealing with time-varying neural activations.

Acknowledgments

This research was supported by grant HFSPO RG0136/1997-B, DFG He1431/8-1, DFG 1431/3-5, BMBF 01ZZ9510 TPA7.

Reprint requests should be sent to Hermann Hinrichs, Department for Neurology II, University of Magdeburg, Leipziger Str. 44, D-39120 Magdeburg, Germany; e-mail: hermann.hinrichs@medizin.uni-magdeburg.de.

The data reported in this experiment have been deposited in National fMRI Data Center (http://www.fmridc.org). The accession number is 2-1000-11158.

REFERENCES

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

Burock, M. A., Buckner, R. L., Woldorff, M. G., Rosen, B. R., & Dale, A. M. (1998). Randomized event-related experimental design allow for extremely rapid presentation rates using functional MRI. *NeuroReport, 9,* 3735–3739.

Burock, M. A., & Dale, A. M. (in press). Estimation and detection of event-related fMRI signals with temporally correlated noise: A statistically efficient and unbiased approach. *Human Brain Mapping*.

Dale, A. M. (1999). Optimal experimental design for event related fMRI. *Human Brain Mapping, 8,* 109–114.

Dale, A. M., & Buckner, R. L. (1997). Selective averaging of rapidly presented individual trials using fMRI. *Human Brain Mapping, 5,* 329–340.

Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. D., & Turner, R. (1998). Event-related fMRI: Characterizing differential responses. *Neuroimage, 7,* 30–40.

Ganis, G., Kutas, M., Schendan, M. E., & Dale, A. M. (1997). DOC: A direct method for temporal overlap correction of ERP and fMRI responses. *Cognitive Neuroscience Fourth Annual Meeting Abstracts, 4,* 42.

Goodyear, B. G., & Menon, R. S. (1998). Effect of luminance contrast on BOLD fMRI response in human primary visual areas. *Journal of Neurophysiology, 79*, 2201–2207.

Hamilton, J. D. (1994). *Time series analysis* (pp. 220–221). Princeton, NJ: Princeton University Press.

Hansen, J. C. (1983). Separation of overlapping waveforms having known temporal distributions. *Journal of Neuroscience Methods, 9,* 127–139.

Heinze, H.-J., Mangun, G. R., Burchert, W., Hinrichs, H., Scholz, M., Münte, T. F., Gös, A., Johannes, S., Scherg, M., Hundeshagen, H., Gazzaniga, M. S., & Hillyard, S. A. (1994). Combined spatial and temporal imaging of spatial selective attention in humans. *Nature, 392,* 543–546.

Josephs, O., Turner, R., & Friston, K. J. (1997). Event-related fMRI. *Human Brain Mapping, 5,* 243–248.

Luck, S. J., Heinze, H. J., Mangun, G. R., & Hillyard, S. A. (1990). Visual event-related potentials index focused attention within bilateral stimulus arrays II. Functional dissociation of P1 and N1 components. *Electroencephalography and Clinical Neurophysiology, 75*, 528–542.

Petry, H. M., Donovan, W. J., Moore, R. K., Dixon, W. B., & Riggs, L. A. (1982). Changes in human visually evoked cortical potential in response to chromatic modulation of a sinusoidal grating. *Vision Research, 22*, 745–755.

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. (1988). *Numerical recipes in C*. Cambridge: Cambridge University Press.

Rosen, B. R., Buckner, R. L., & Dale A. M. (1998). Event-related functional MRI: Past, present, and future. *Proceedings of the National Academy of Sciences U.S.A., 95*, 773–780.

Talairach, J., & Tournoux, P. (1988). *A co-planar stereotaxic atlas of a human brain*. Stuttgart: Thieme.

Ward, R., Duncan, J., & Shapiro, K. (1996). The slow time-course of visual attention. *Cognitive Psychology, 30,* 79–109.

Weisskoff, R. M., Baker, J., Belliveau, J., Davis, T. L., Kwong, K. K., Cohen, M. S., & Rosen, B. R. (1993). Power spectrum analysis of functionally-weighted MR data: What's in the noise? [Abstract] *Proceedings of the Society for Magnetic Resonance, 1,* 7.

Woldorff, M. G. (1993). Distortion of ERP averages due to overlap from temporally adjacent ERPs: Analysis and correction. *Psychophysiology, 30,* 98–119.

Woldorff, M. G., & Hillyard, S. A. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography and Clinical Neurophysiology, 79,* 170–191.

Xiong, J., Gao, J. H., Lancaster, J. L., & Fox, P. T. (1996). Assessment and optimization of functional MRI analyses. *Human Brain Mapping, 4,* 153–167.

Zarahn, E., Aguirre, G. K., & D'Esposito, M. (1997). Empirical analysis of BOLD fMRI statistics: Spatially unsmoothed data collected under null-hypothesis conditions. *Neuroimage, 5,* 179–197.