Henson R.N. (2015) Design Efficiency. In: Arthur W. Toga, editor. *Brain Mapping: An Encyclopedic Reference*, vol. 1, pp. 489-494. Academic Press: Elsevier.

Design Efficiency

RN Henson, MRC Cognition and Brain Sciences Unit, Cambridge, UK

Abbreviations

AR(p)	Autoregressive model of order p
BOLD	Blood oxygenation level-dependent (signal normally measured with fMRI)
DCT	Discrete cosine transform
FIR	Finite impulse response (basis set)
fMRI	Functional magnetic resonance imaging
GLM	General linear model
HRF	Hemodynamic response function
ReML	Restricted maximum likelihood
SOA	Stimulus-onset asynchrony (SOA_{min} = minimal SOA)
TR	Interscan interval (repetition time)

Formal Definition of Efficiency

The general linear model (GLM) normally used for mass-univariate statistical analysis of functional magnetic resonance imaging (fMRI) data can be written for a single voxel as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim N(0, \sigma^2 \mathbf{C}_e) \qquad [1]$$

where $\mathbf{y}$ is an $N \times 1$ column vector of the data time series sampled every TR for N scans, $\mathbf{X}$ is an $N \times P$ *design matrix* in which the P columns are regressors for the time series of predicted experimental effects, $\boldsymbol{\beta}$ is a $P \times 1$ column vector of parameters for each regressor in $\mathbf{X}$ (whose values are estimated when fitting the model to the data), and $\boldsymbol{\epsilon}$ is $N \times 1$ vector of residual errors. The second expression in eqn [1] denotes that the residuals come from a zero-mean, multivariate normal (Gaussian) distribution with covariance $\mathbf{C}_e$. Normally, the residuals are assumed to be drawn independently from the same distribution (*white* residuals), or if not, then the data and model are filtered, or *prewhitened*, by an estimate of the error covariance (see later). This means that $\mathbf{C}_e = \sigma^2 \mathbf{I}$, corresponding to an $N \times N$ identity matrix ($\mathbf{I}$) scaled by a single variance term σ^2.

Assuming white residuals, the parameters can be estimated by minimizing the sum of squares of the residuals, to give the so-called *ordinary least squares* (OLS) estimates, $\hat{\boldsymbol{\beta}}$. The planned comparisons we want to test with our experiment are a linear combination of these parameter estimates, specified by a $1 \times P$ *contrast vector*, $\mathbf{c}$. For example, $\mathbf{c} = [1 \quad -1]$ would test whether the parameter estimate for the first of two regressors is greater than the second. Significance can be assessed by a *T*-statistic, defined by

$$T(\text{df}) = \frac{\mathbf{c}\hat{\boldsymbol{\beta}}}{\sqrt{\mathbf{c}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{c}^T\hat{\sigma}^2}} \qquad [2]$$

where $\hat{\sigma}^2$ is the error variance estimated by $\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}/\text{df}$ (where^T denotes the transpose of a matrix) and the degrees of freedom, df, are defined by $N - \text{rank}(\mathbf{X})$. The probability, p, of getting a value of T or greater under the null hypothesis that $\mathbf{c}\hat{\boldsymbol{\beta}} = 0$, given the df, can then be calculated from Student's *T*-distribution, and the null hypothesis rejected if, for example, $p < 0.05$.

We are now in the position to define the *efficiency* of a contrast, e, as

$$e = 1\Big/\Big(\mathbf{c}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{c}^T\Big) \qquad [3]$$

which can be seen as inversely related to the denominator of the *T*-statistic in eqn [2]. Thus, if we increase e, we also increase T. (For multiple contrasts, where $\mathbf{c}$ is an $M \times P$ matrix of M contrasts, such as an *F*-contrast, we can define the average efficiency as $1/\text{trace}\{\mathbf{c}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{c}^T\}$.)

Note that the scaling of e is arbitrary (depending on the scaling of the contrast, scaling of regressors, and number of scans), so the precise relationship between e and T is best assumed only to be monotonic. Note also that this statement assumes that the estimate of the error variance ($\hat{\sigma}^2$) is independent of the design ($\mathbf{X}$), which may not always be true (see later). Given these assumptions, and that the contrasts are specified a priori, then to maximize the efficiency of our design, we simply need to vary $\mathbf{X}$. We now consider how $\mathbf{X}$ is defined for fMRI.

HRF Convolution

We can start by assuming that stimuli elicit brief bursts of neural activity, or *events*, which are modeled by delta functions every time a stimulus is presented. Then, for the jth of N_j event types (conditions), the neural activity over time, or *neural time course*, $u_j(t)$, can be expressed as

$$u_j(t) = \sum_{i=1}^{i=N_i(j)} \delta(t - T_{ji})$$

where T_{ji} is a vector of $i = 1 \ldots N_i(j)$ onset times and δ is the Dirac delta function. With fMRI, we do not measure neural activity directly, but rather the delayed and dispersed BOLD impulse response, $b(\tau)$, where τ indexes poststimulus time (e.g., from 0 to 30 s). Given that $b(\tau)$ may vary across voxels (and individuals), it can be modeled by linear combination of N_k hemodynamic response functions (HRFs), $h_k(\tau)$:

$$b_j(\tau) = \sum_{k=1}^{k=N_k} \beta_{kj} h_k(\tau)$$

where β_{kj} are the parameters to be estimated for each HRF and condition (and voxel).

Assuming that BOLD responses summate linearly (though see later), the predicted BOLD time course over the experiment, $x(t)$, can then be expressed as the convolution of the neural time courses by the HRFs:

$$x(t) = \sum_{j=1}^{j=N_j} u_j(t) \otimes b(\tau) = \sum_{j=1}^{j=N_j} \sum_{k=1}^{k=N_k} \sum_{i=1}^{i=N_i(j)} \beta_{kj} h_k(t - T_{ji}) \quad [4]$$

resulting in a linearly separable equation that can be represented by a design matrix $\mathbf{X}$ with $P = N_j N_k$ columns.

At one extreme, we can assumed a fixed shape for the BOLD response by using a single *canonical* HRF (i.e., $N_k = 1$). At the other extreme, we can make no assumptions about the shape of the BOLD response (up to a certain frequency limit) by using a so-called *finite impulse response* (FIR) set (see Figure 1; for multiple basis functions, the contrasts become $\mathbf{c} \otimes \mathbf{I}_{N_k}$, where $\mathbf{c}$ is a contrast across the N_j event types and $\mathbf{I}_{N_k}$ is an $N_k \times N_k$ identity matrix for the N_k basis functions). Normally, one is only interested in the magnitude of a BOLD response, in which case a single canonical HRF is sufficient to estimate efficiency a priori (by assuming that a canonical HRF is a sufficient approximation on average across voxels and individuals). If however one is interested in estimating the shape

of the BOLD impulse response, then a more general set such as an FIR is necessary. Using a canonical HRF would correspond to what Liu, Frank, Wong, and Buxton (2001) called *detection power*, while using an FIR would correspond to what they called *estimation efficiency*. This is important because the choice of HRF affects the optimal experimental design (see later).

Filtering

So far, we have considered definition of the signal, $x(t)$, but the other factor that affects the T-statistic in eqn [2] is the noise variance, $\hat{\sigma}^2$. fMRI is known to have a preponderance of low-frequency noise, caused, for example, by scanner drift and by biorhythms (e.g., pulse and respiration) that are aliased by slower sample rates (1/TR). A common strategy therefore is to high-pass filter the data. An example matrix, $\mathbf{F}$, for implementing high-pass filtering within the GLM using a discrete cosine transform (DCT) set is shown in Figure 1. The reduction in noise will improve sensitivity, as long as the filtering does not remove excessive signal too. Heuristics suggest that an
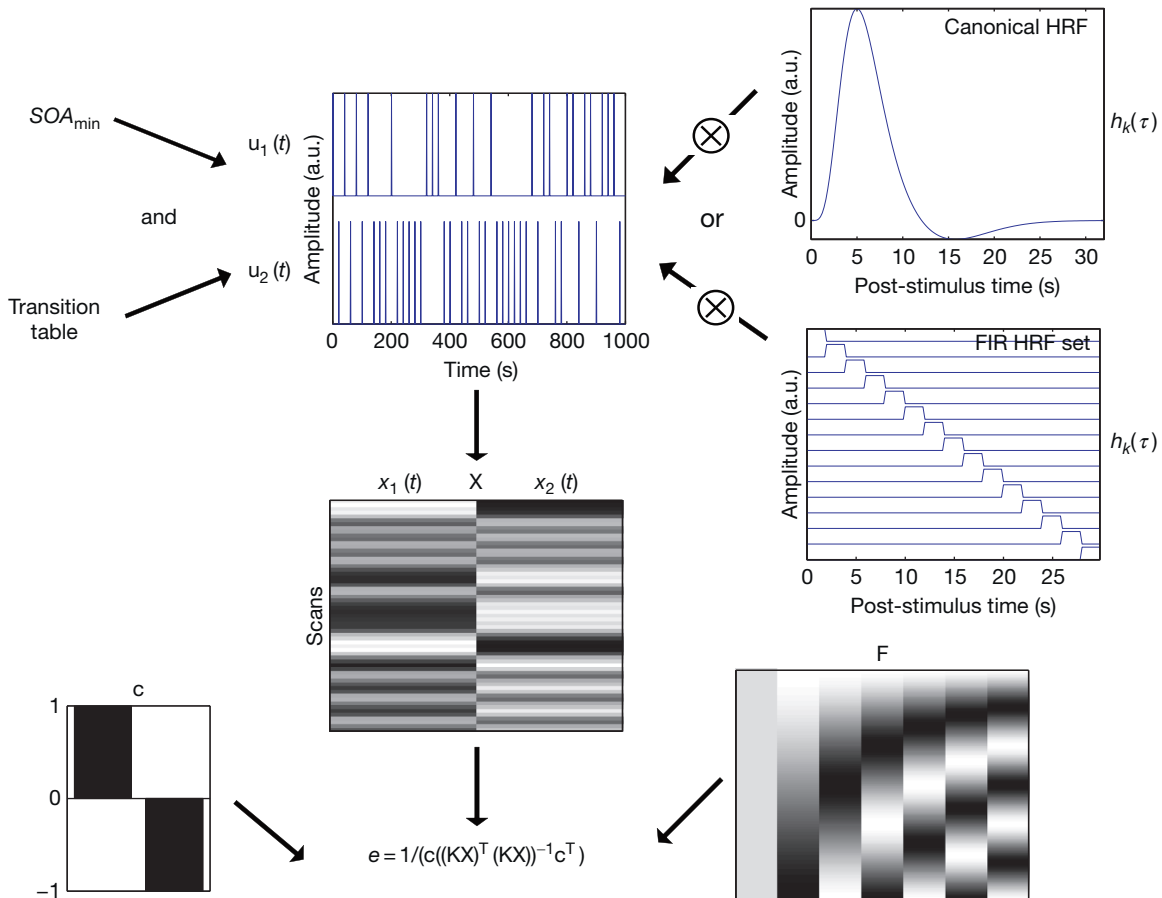
Figure 1 Ingredients for efficiency: the minimal SOA, SOA_{min}, and stimulus *transition table* determine the neural time course, $u_j(t)$, which is convolved with HRFs of poststimulus time, $h_k(\tau)$, to create the design matrix, $\mathbf{X}$. This, together with an a priori contrast, $\mathbf{c}$, and any (high-pass) filter matrix, $\mathbf{K}$ (here generated by $\mathbf{K} = \mathbf{I}_N - \mathbf{FF}^{-1}$), then determines the efficiency, e.

approximate inflection in the noise power spectrum typically occurs at around 1/120 s, which is why it is inadvisable to have designs with changes in signal slower than this (e.g., alternating blocks of more than 60 s; see later).

High-pass filtering also helps render the noise white, that is, constant across frequencies, though there is often still temporal autocorrelation (color) in the residuals. A common strategy to deal with this is to hyperparameterize the error covariance matrix $\mathbf{C}_e$ using an AR(p) model, which can be estimated using ReML (Friston et al., 2002). Once estimated, $\mathbf{C}_e$ can be inverted in order to prewhiten the data and model, which therefore also affects efficiency. Mathematically, both high-pass filtering and prewhitening can be implemented by multiplying data and model by a single filter matrix, $\mathbf{K}$, such that efficiency becomes

$$1 \left/ \left(\mathbf{c} \left((\mathbf{KX})^{\mathrm{T}} (\mathbf{KX}) \right)^{-1} \mathbf{c}^{\mathrm{T}} \right) \right. \qquad [5]$$

Parameterizing Experimental Designs

For events, the neural time course is determined by T_{ji} in eqn [4], which itself can be captured by two parameters: (1) the minimal time between events, or *minimal stimulus-onset asynchrony*, $\mathrm{SOA}_{\mathrm{min}}$, and (2) a *transition table*, which can be defined by two matrices, an $N_P \times N_H$ matrix, $TM^{(p)}$, describing the history of the *previous* N_H event types, of which there are N_P possible sequences (in the extreme case $N_P = N_J^{N_H}$, though it can be smaller), and an $N_P \times N_J$ matrix, $TM^{(n)}$, describing the probability of the next event being one of the N_J event types, given each of those possible previous sequences. So for a fully randomized design, where the probability of each event is equal and independent of previous events ($N_H = 1$), $TM^{(p)}$ would be $N_J \times 1$ matrix and $TM^{(n)}$ would be $N_J \times N_J$ matrix with values of $1/N_J$ for each event type. So for $j = 1, 2$ event types,

$$TM^{(p)} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} TM^{(n)} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

This implies that there is an equal chance of event type 1 being followed by event type 1 as there is for it being followed by event type 2 and likewise for what follows event type 2. Specifying a design in terms of probabilistic transition matrices allows one to treat the design matrix as a random variable and derive the expected efficiency by averaging over all possible design matrices (see Friston, Zarahn, Josephs, Henson, & Dale, 1999, for details). In other words, one can express design efficiency in terms of the probabilistic contingencies entailed by the design matrix.

Randomized Designs

For a randomized design with two event types, we can plot the efficiency against $\mathrm{SOA}_{\mathrm{min}}$ for each of 2 contrasts, $\mathbf{c} = \begin{bmatrix} 1 & -1 \end{bmatrix}$, the *differential effect* of event types 1 and 2, and $\mathbf{c} = \begin{bmatrix} 1 & 1 \end{bmatrix}$, the *common effect* of event types 1 and 2 versus the interstimulus

baseline. For a canonical HRF, the efficiency of these two contrasts is plotted against $\mathrm{SOA}_{\mathrm{min}}$ in Figure 2(a). As can be seen, the optimal SOA for the common effect is around 18 s, whereas the optimal SOA for the differential effect increases exponentially as $\mathrm{SOA}_{\mathrm{min}}$ decreases.

The basic reason for these results is that higher efficiency corresponds to greater variability of the signal over time (where the signal is a function of the contrast and regressors). At short SOAs, the sluggish (low-pass) nature of the HRF means that when we do not distinguish the two event types (by using a contrast for their common effect), the BOLD responses for successive events summate to give a small oscillation around a raised baseline (leftmost inset in Figure 2(a)), that is, low signal variance. The random ordering of the two event types means that for the differential effect, however, there is a large variance in signal. At longer SOAs, there is time for the BOLD response to return to baseline between events, so signal variance (efficiency) is increased for the common effect, but the variance for the differential effect decreases relative to short SOAs (rightmost inset in Figure 2(a)). For further explanations of this behavior of efficiency, for example, in terms of signal processing or correlations between regressors, see http://imaging.mrc-cbu.cam.ac.uk/imaging/DesignEfficiency. The important point of Figure 2(a) is that if one is interested in the difference between two randomly ordered event types, then a shorter SOA is generally better, though the price one pays is reduced efficiency to detect the common effect of both versus baseline. In reality though, this increasing efficiency with decreasing $\mathrm{SOA}_{\mathrm{min}}$ (for a differential effect) cannot occur indefinitely, because at some point, there will be saturation of the neural and/or hemodynamic response to stimuli that are too close together in time, an example of nonlinear behavior that is ignored under the linear superposition assumptions made so far (see later).

One can improve efficiency for the common effect versus baseline at short SOAs by ensuring that the probabilities in $TM^{(n)}$ do not sum to 1 (across columns). This means that for some $\mathrm{SOA}_{\mathrm{min}}$, no event occurs (sometimes called a *null-event* or *fixation trial*, Dale, 1999). This effectively produces a stochastic distribution of SOAs, with an exponentially decreasing probability of each SOA. The efficiency for the common effect now also increases as $\mathrm{SOA}_{\mathrm{min}}$ decreases, at a cost to the efficiency of the differential effect (Figure 2(b)). These types of design are suitable for an FIR basis set (insets in Figure 2(b)) because a stochastic distribution of SOAs allows the BOLD response shape to be estimated, particularly with short $\mathrm{SOA}_{\mathrm{min}}$. (Note that efficiency is not directly comparable across canonical and FIR HRFs, since it depends on the scaling of the basis functions, and an FIR will also entail a reduction in the df, which will affect the final T-statistic in eqn [2].)

Blocked Designs

Events of the same type can be blocked into short sequences, which can increase the detection power relative to a randomized design. For a blocked design with two event types, there would be N_H events per block; for example, for blocks of three events,
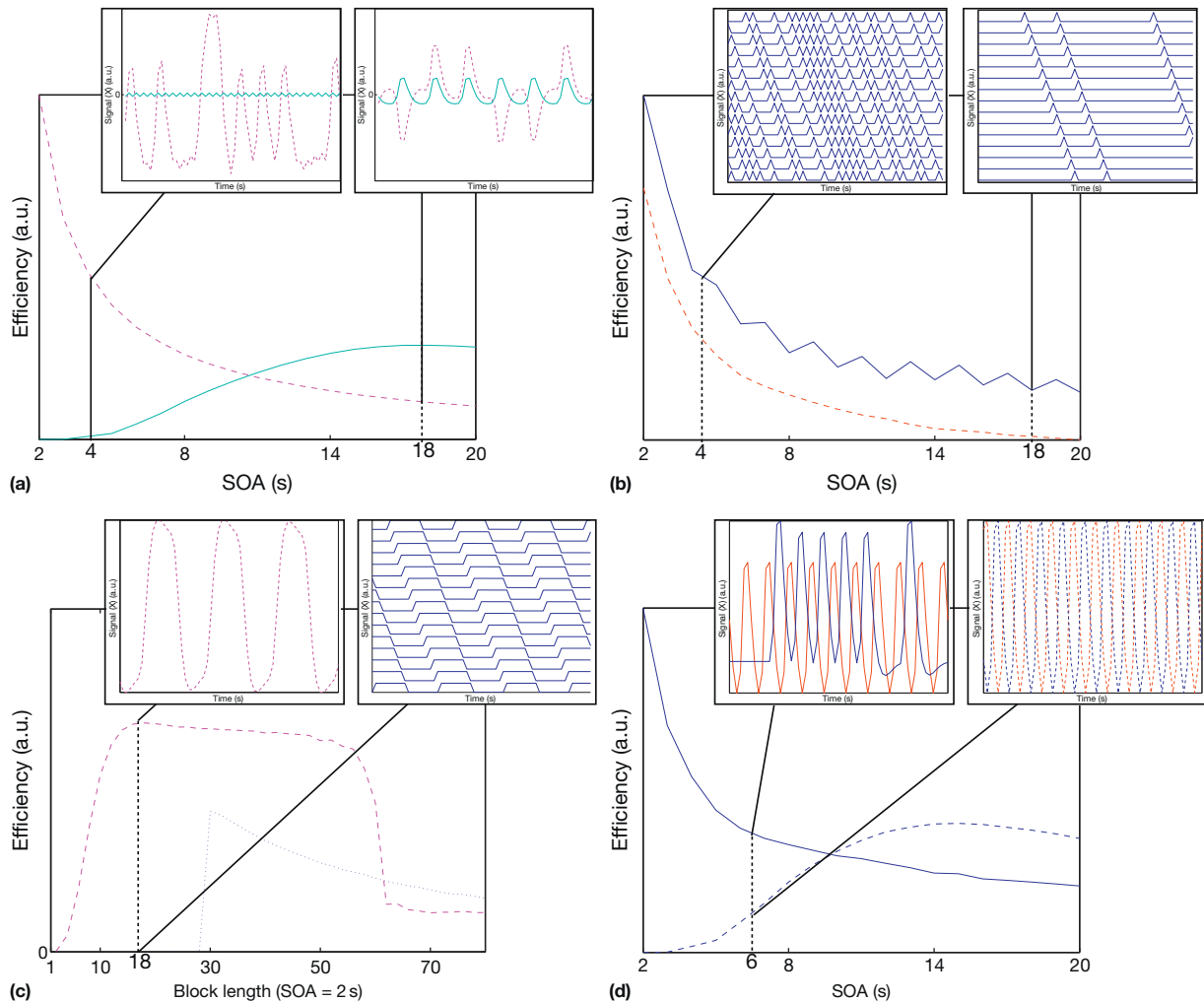
Figure 2   Efficiency for various possible contrasts and designs. (a) Efficiency (arbitrary units) as a function of SOA_{min} in a randomized design using a canonical HRF for a differential $[1 \ -1]$ contrast between two event types (dashed magenta line) and the common $[1 \ 1]$ contrast versus baseline (solid cyan line). Insets are sections of corresponding contrasts of regressors (predicted signal) for $SOA_{min} = 4$ (left) and $SOA_{min} = 18$ (right). Efficiency for the differential effect is higher at short SOAs owing to the greater signal variance caused by the random ordering of event types. (b) Similar to (a), but now including null events with probability 1/3 and an FIR basis set (dashed red = differential effect; solid blue = common effect). The stochastic distribution of SOAs caused by null events increases efficiency for common effect versus baseline even at short SOA_{min}. (c) Efficiency as function of block length for a differential contrast between two event types in a blocked design (and high-pass filter cutoff of 120 s). The dashed magenta line and left inset correspond to a canonical HRF; the dashed cyan line and right inset correspond to an FIR basis set. Maximal efficiency with a canonical HRF arises for a block length of 18 s; for an FIR basis set, blocks shorter than the FIR duration are inefficient, owing to linear dependence between the basis functions. (d) Efficiency for the unique effect of the second of two event types ($[0 \ 1]$ contrast) using a canonical HRF in a design where the second event type can only follow the first event type. The dashed blue line and rightmost inset show an alternating design in which the second event type always follows the first; the solid blue line and left inset show a design in which the second event type follows the first 50% of the time (the red dashed line in the insets corresponds to the regressor for the first event type). For SOA_{min} below approximately 10 s (e.g., 6 s), the 50% design is more efficient (despite fewer events of the second type in total), because it decorrelates the two regressors. In all panels, 2000 scans with TR = 2 s were simulated, with the first 30 s discarded to remove transient effects.

$$TM^{(p)} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix} \quad TM^{(n)} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}$$

The magenta dashed line in Figure 2(c) shows efficiency for the differential effect between two event types as a function of block length using a canonical HRF when SOA_{min} is 2 s and the high-pass cutoff is 120 s. Short blocks (in the extreme case,

alternating event types when block length is 1) have low efficiency, for the same reason that the main effect is inefficient at short SOAs: any variance in neural activity is smoothed out by the HRF. As in Figure 2(a), efficiency is maximal for block lengths around 18 s (since 1/18 Hz is close to the highest passband of the canonical HRF filter; Josephs & Henson, 1999), but for block lengths of 30 s or more, efficiency plummets again because of the high-pass filter: for such long blocks, most of the signal variance (particularly that at the fundamental frequency of the block alternation) is low enough to be removed by the high-

pass filter, which is appropriate since it is likely to be masked by fMRI noise anyway. In short, one does not want fMRI designs where the signal changes too fast (since it will be attenuated by the sluggish BOLD response) or where it changes too slow (since it will be removed by any high-pass filtering or swamped by noise).

The blue dotted line in **Figure 2(c)** shows efficiency as a function of block length when using an FIR basis set rather than canonical HRF. In this case, efficiency is low, at least until the block length exceeds the duration of the assumed BOLD response (30 s here). This is because of the high covariance between regressors for each basis function (rightmost inset in **Figure 2(c)**), as explained later. This illustrates that, while blocked designs are efficient for estimating the amplitude of a known BOLD response, they are not efficient for estimating the unknown shape of a BOLD response (Liu et al., 2001).

Unique Effects

In addition to the differential and common effects for two event types, one might be interested in the unique effect of each event type, having adjusted for effects of the other event type, which corresponds to contrasts of $[1 \quad 0]$ or $[0 \quad 1]$. A common example is when one type of event must always follow another type, such as a motor act that is contingent on a stimulus (e.g., a working memory trial). The dashed line in **Figure 2(d)** shows that, for SOAs below 9 s, efficiency is low for a $[0 \quad 1]$ contrast in such an alternating design. To improve efficiency, one needs to reduce the correlation between the two event types (since the term $\mathbf{X}^T\mathbf{X}$ in eqn [3] relates to the covariance of the regressors in the design matrix). One way to do this is to randomize the SOA between stimuli and motor acts; another is to only require a motor act on a random fraction, say 50%, of trials, as shown in the solid line in **Figure 2(d)**. Although this entails fewer motor events in total, the efficiency for separating the BOLD response to the motor act from that to the stimulus is increased at short SOAs.

A similar issue arises in so-called *mixed* designs, in which one wishes to separate a sustained, or *state*, effect from a transient, or *item*, effect (Chawla, Rees, & Friston, 1999). This requires blocks of events, in which the SOA is varied within each block so as to reduce the correlation between the (epoch) regressor modeling the state effect and the (event) regressor modeling the item effect. The downside of such designs is that this requirement to decorrelate the two regressors, in order to estimate both unique effects reasonably efficiently, considerably reduces the efficiency for estimating either effect alone, relative to designs with only epochs, or only events.

Optimizing Designs

The examples in **Figure 2** represent just a subspace of possible designs, chosen to help illustrate some of the properties of eqn [5]. Other formal explorations of design space can be found in, for example, Dale (1999), Josephs and Henson (1999), Friston et al. (1999), and Hagberg, Zito, Patria, and Sanes (2001). There are automated ways of maximizing efficiency by searching through possible designs, using, for example, genetic algorithms (Wager & Nichols, 2003). Additional constraints are often needed however, such as limits on runs of the same event type; otherwise, an optimization scheme is likely to converge on a blocked design, which is always most efficient (for detection power) from the fMRI perspective, but may not be appropriate from the psychological perspective (e.g., if the presence of structure in the sequence of events affects brain activity). An interesting class of pseudorandomized design that has optimal estimation efficiency is an *m-sequence* (Buracas & Boynton, 2002). This is a deterministic sequence that presents all combinations of event histories up to $N_H = m$ (i.e., has a large, deterministic transition matrix) but that is nonetheless effectively unpredictable to participants. Such sequences have been computed for $N_J = 2$, 3, and 5 event types but require a fixed number ($N_J^{N_H} - 1$) of events in total.

Nonlinearities and Assumptions

The main assumption made in the examples mentioned earlier is that the brain's response to successive events is a linear superposition of responses that would occur to each event on its own. Of course, this is unlikely in practice, particularly for short SOAs, where there is likely to be saturation of neural and/or hemodynamic responses. Such saturation has been demonstrated empirically and can be modeled with the GLM by using Volterra kernels (Friston, Josephs, Rees, & Turner, 1998). Once fit, these kernels can be used to adjust predictions for efficiency. For kernels fit to auditory cortex responses to auditory stimuli of varying SOAs, the negative impact of saturation on efficiency still only became appreciable for SOAs below 2 s: that is, there was still an advantage of short SOAs down to 2 s in randomized designs.

A second assumption is that changing the design ($\mathbf{X}$ in eqn [5]) does not affect the estimation of the error ($\hat{\sigma}^2$ in eqn [2] or even $\mathbf{K}$ in eqn [5]). Different SOAs (even across trials within a design) may entail differences in HRF shape, or different nonlinearities, resulting in inaccurate model fits and therefore different residuals. Indeed, when comparing blocked and randomized designs, differences in the error estimate have been shown empirically (Mechelli, Price, Henson, & Friston, 2003). Nonetheless, without such a priori knowledge about the noise, one can only rely on the general heuristics about maximizing the signal outlined earlier.

Summary

Efficiency is a well-defined mathematical property of the GLM, and under the linear assumptions of a convolution model for the BOLD response, efficiency can be optimized for a priori contrasts of the conditions of an fMRI experiment by selecting the optimal SOA_{min} and stimulus transition table.

Acknowledgments

See also: **INTRODUCTION TO METHODS AND MODELING**:
Contrasts and Inferences; Convolution Models for FMRI; The General
Linear Model.

References

Buracas, G. T., & Boynton, G. M. (2002). Efficient design of event-related fMRI experiments using M-sequences. *NeuroImage, 16*, 801–813.

Chawla, D., Rees, G., & Friston, K. J. (1999). The physiological basis of attentional modulation in extrastriate visual areas. *Nature Neuroscience, 2*, 671–676.

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

Friston, K. J., Josephs, O., Rees, G., & Turner, R. (1998). Non-linear event-related responses in fMRI. *Magnetic Resonance in Medicine, 39*, 41–52.

Friston, K. J., Zarahn, E., Josephs, O., Henson, R. N., & Dale, A. M. (1999). Stochastic designs in event-related fMRI. *NeuroImage, 10*, 607–619.

Friston, K. J., Glaser, D. E., Henson, R. N., Kiebel, S., Phillips, C., & Ashburner, J. (2002). Classical and Bayesian inference in neuroimaging: Applications. *Neuroimage, 16*, 484–512.

Hagberg, G. E., Zito, G., Patria, F., & Sanes, J. N. (2001). Improved detection of event-related functional MRI signals using probability functions. *Neuroimage, 14*, 1193–1205.

Josephs, O., & Henson, R. N. (1999). Event-related fMRI: Modelling, inference and optimisation. *Philosophical Transactions of the Royal Society, London, 354*, 1215–1228.

Liu, T. T., Frank, L. R., Wong, E. C., & Buxton, R. B. (2001). Detection power, estimation efficiency, and predictability in event-related fMRI. *Neuroimage, 13*, 759–773.

Mechelli, A., Price, C. J., Henson, R. N., & Friston, K. J. (2003). Estimating efficiency a priori: A comparison of blocked and randomised designs. *Neuroimage, 18*, 798–805.

Wager, T. D., & Nichols, T. E. (2003). Optimization of experimental design in fMRI: A general framework using a genetic algorithm. *Neuroimage, 18*, 293–309.

Relevant Websites

http://www.cabiatl.com/CABI/resources/fmrisim/ – Tool for simulating fMRI designs.

http://www.mrc-cbu.cam.ac.uk/wp-content/uploads/2013/09/fMRI_GLM_efficiency.m – Matlab code used to calculate efficiency in examples here.

http://imaging.mrc-cbu.cam.ac.uk/imaging/DesignEfficiency – General advice about how to optimise an fMRI experiment.

http://psych.colorado.edu/~tor/Software/genetic_algorithms.html – Genetic algorithm for optimising fMRI designs.

http://surfer.nmr.mgh.harvard.edu/optseq/ – Tool for optimising randomised designs.