A test of time-resolved functional MRI with sub-second event durations

Alvin P.H. Wong[a,b], Esther X.W. Wu[b,c], Baxter P. Rogers[d], & Christopher L. Asplund[a,b,c,e,f,*]

[a] Department of Psychology, Faculty of Arts and Social Sciences, National University of Singapore, 5 Arts Link, Singapore 117570

[b] N.1 Institute for Health, National University of Singapore, 28 Medical Drive, #05-COR, Singapore 117456

[c] Division of Social Sciences, Yale-NUS College, National University of Singapore, 16 College Ave West, Singapore 138527

[d] Department of Radiology and Radiological Sciences, Vanderbilt University, 1161 21st Avenue South, Nashville, TN, USA 37232

[e] Centre for Sleep and Cognition, Yong Loo Lin School of Medicine, 14 Medical Drive, #B1-01, Singapore 117599

[f] Neuroscience and Behavioural Disorders, Duke-NUS Medical School, 8 College Road, Singapore 169857

*To whom correspondence should be sent. Address: Yale-NUS College, 28 College Avenue West #01-501, Singapore 138533. Tel: +65 6601 3327. Email: chris.asplund@yale-nus.edu.sg.

Abstract

Functional neuroimaging (fMRI) signals are famously sluggish, unfolding over several seconds after even brief sensory (and neural) events. Relative timing differences between different evoking conditions, however, can distinguish between changes in the duration or intensity of neural activity (Henson et al., 2002). Specifically, increases in stimulus duration cause a delayed response peak with a greater amplitude, whereas increases in stimulus intensity affect the peak magnitude alone. These relationships have been empirically demonstrated using stimulus durations of several seconds (Friston, 2005), but many processes relevant to cognitive psychology occur on sub-second timescales. Here we tested whether the predicted relationship between stimulus properties and the hemodynamic response held for sub-second stimuli and sub-second changes in stimulus duration. In Experiment 1 (n=15), we presented visual and auditory stimuli at three durations (100, 300, and 900 ms) and three intensities in a slow event-related design with a rapid TR (625 ms). In Experiment 2 (n=14), the stimulus durations were extended (1000, 1200, and 1800 ms) to increase the evoked signal. By fitting the observed fMRI signals to parameterised hemodynamic response functions, we found that changes in stimulus duration caused differences in peak latency and magnitude, whereas changes in stimulus intensity affected only peak magnitude. Significant effects were found for duration differences as small as 200 ms, with clearer results when a sufficiently large base signal was evoked and when a bootstrap resampling technique was used. Our results support the validity of existing models of the hemodynamic response. They also support the use of time-resolved fMRI for mental chronometry in cognitive neuroscience, thereby allowing researchers to draw additional inferences about brief mental events.

Introduction

Functional magnetic resonance imaging (fMRI) signals unfold over several seconds, even when caused by brief mental events. An increase in either neural activation intensity or duration is primarily reflected as an increase in the sluggish BOLD (blood oxygen-level dependent) response's amplitude. As such, BOLD amplitude is the key dependent measure in most neuroimaging studies, even when the manipulations are hypothesized to affect the duration of neurocognitive processes (Henson et al., 2002; Poldrack, 2015). Nevertheless, the BOLD signal can preserve information about the timing of mental events as well. Although duration and intensity changes have similar effects on the measured signal, relative timing differences can distinguish their cause (Henson et al., 2002). Specifically, an increase in stimulus duration leads to both an increase in the hemodynamic response's peak latency (time to peak after the onset of the evoking stimulus) and peak magnitude. Conversely, increases in stimulus intensity increase peak magnitude alone (Friston, 2005).

Many studies have relied on this dissociation to test psychological theories. The theories typically involve bottlenecks of information processing, for which patterns of increased processing durations are the critical diagnostic feature. As such, they rely on the logic of mental chronometry, in which inferences about the properties of psychological events are based on changes in temporal measures, such as reaction time (Lo & Andrews, 2015). Peak latency represents another such measure. Peak latency shifts have been used to argue for, and to establish neural correlates of, bottlenecks in multitasking (Dux et al., 2006, 2009), attentional and response selection (Asplund et al., 2010; Marti et al., 2012; Tombu et al., 2011), and language comprehension (Vagharchakian et al., 2012). In each case, the behavioral and peak latency timing differences across conditions were small, typically on the order of 200-1000 ms.

The small magnitude of these timing effects is not unexpected, but it does go beyond the empirical evidence for the relationship between stimulus properties and hemodynamic response features. Distinctions between the duration and intensity of neural activity are reflected well in the BOLD signal with large timing differences (2 versus 6 seconds, for example; Friston, 2005), but whether smaller differences yield the same differences is currently unknown. Prior relevant studies either used long stimulus durations or stimulus duration differences (e.g. >1 s) (Liu & Gao, 2000; Vazquez & Noll, 1998), employed brief stimulus duration manipulations without varying stimulus intensity (Lewis et al., 2018), or did not compare effects of intensity and duration to determine whether they could be reliably distinguished (Vazquez & Noll, 1998; Yeşilyurt et al., 2008). As such, the approach underlying the aforementioned bottleneck studies has not been empirically tested within the relevant range of stimulus features.

Even if the predicted relationships between stimulus properties and hemodynamic features hold for brief stimuli, it is unclear whether fMRI has the temporal resolution to reliably detect the subtle timing effects. Although the practical temporal effect limit is not known, timing differences as low as 28 ms have been reliably detected (Katwal et al., 2013; Rogers et al., 2010; Wang et al., 2017) for shifts in the hemodynamic response. Onset differences of under 100 ms have been reliably detected in other contexts (Setzer et al., 2022). In both cases, characterization of the hemodynamic response appeared to be improved with a high TR (repetition time), which allowed for more overall samples as well as more fine-grained estimation of the hemodynamic function's shape (Dowdle et al., 2021; Polimeni & Lewis, 2021). Timing differences on the order of hundreds of milliseconds may be on the edge of standard fMRI paradigms' resolution, but newer MR sequences and good design choices may allow them to be reliably detected.

Another approach that may assist with detecting subtle timing effects is the use of parameter estimates from fitted HRFs. In this approach, a parameterized model is fitted to the data for each participant and condition, effectively filtering the data to reduce noise (Friston, 2005). The derived parameters are then used for statistical analyses and comparisons. Parameters such as peak magnitude, peak latency, response latency and peak dispersion have been shown to constitute meaningful indicators of both within-subject and between-group differences (Wager et al., 2005). HRF fitting approaches can also be combined with other techniques to further increase the signal for temporal metrics. For example, researchers using ERPs (evoked response potentials) have used resampling techniques, such as bootstrap resampling (Efron & Gong, 1983), to obtain more reliable peak timing estimates (Bledowski, 2004; Liesefeld, 2018; Miller et al., 1998).

In the present study, we sought to characterize and compare the hemodynamic responses to visual and auditory stimuli that varied in their durations and intensities. In Experiment 1 (n=15), we presented stimuli at three durations (100, 300, and 900 ms) and three intensities (different contrast or volume). To best capture the temporal features of the evoked responses, we used a slow event-related design with a rapid TR (625 ms) and then derived key parameters (e.g. peak latency and magnitude) from fitted double-Gamma HRFs, both with and without bootstrap resampling. We hypothesized that peak magnitude would increase with increasing stimulus durations or intensities, whereas only peak latency would increase with increasing stimulus durations. In Experiment 2 (n=14), the stimulus durations were extended (1000, 1200, and 1800 ms) to increase the evoked signal. This change also made our experiment more relevant to studies with overlapping evoked signals (Serences, 2004). We hypothesized the same pattern of results as before, with peak latency uniquely sensitive to stimulus duration changes.

Experiment 1

In this first experiment, we sought to characterize the effects of parametrically-modulated visual and auditory stimuli on evoked hemodynamic responses. Stimulus duration and intensity each had three levels, creating a 3x3 design. We hypothesized that, provided noise did not overwhelm our signals, intensity and duration changes would affect observed BOLD magnitude, but only duration changes would affect peak timing (Henson et al., 2002). Gamma-based hemodynamic response functions were used for formalizing predictions about expected responses and for fitting the empirical data. Parameters derived from these fits were used for standard statistical analyses as well as a bootstrap-based approach inspired by EEG timecourse analyses (Bledowski, 2004; Efron & Gong, 1983; Liesefeld, 2018).

It is worth mentioning that Experiment 1's results were suggestive of the hypothesized effects, but they were also inconclusive. The results informed the design of Experiment 2, which provided stronger evidence to address—and support—our primary hypotheses.

Materials and Methods

Experimental procedures

Sixteen members (9 females, 14 right-handed, mean (±SD) age = 23±2.42 years) of the National University of Singapore (NUS) community were recruited for the study, which took place at the Yong Loo Lin School of Medicine's Clinical Imaging Research Centre (CIRC). Written informed consent was sought and obtained from all participants (for Experiments 1 and 2), in accordance with a protocol approved by the NUS Institutional Review Board. Participants were instructed to avoid caffeine intake in the 24 hours preceding their scanning session. They were paid 40 SGD for their participation.

Participants completed a 70-minute slow event-related main task, conducted across nine 390s runs. Each run began with a 20s fixation period and concluded with a 10s fixation

period, with 18 visual trials and 18 auditory trials presented in the interim. Visual trials began approximately every 20th second, with the precise onset jittered from this time by up to 625ms. This jitter reduced participant expectation of event onsets (including based on rhythmic scanner noise) and potentially increased our temporal resolution through subsampling. Auditory trials were presented with similar timing and jitter, save they began 30s into each run. Visual and auditory trials were thus interleaved, with an approximately 10s fixation period between stimuli.

The experiment was presented on a MacBook Air (OS 10.12.1) running PsychoPy v1.83.01 (Pierce, 2007) and connected to a 32-inch LCD monitor (NordicNeuroLab, Bergen, Norway) running at 60 Hz. Each visual trial consisted of a circular contrast-reversing checkerboard (flicker rate: 10Hz) that was 20.7 degrees wide. Participants viewed the stimuli in a supine position via an angled mirror located 12 cm above the eyes. In turn, the distance between this mirror and the LCD monitor was 179 cm. There were 9 possible trial conditions defined by the conjunction of three stimulus durations (100ms, 300ms, and 900ms) and three stimulus intensities (1% contrast, 10% contrast, and 100% contrast). These contrast levels were chosen because increases in visual stimulus intensity by powers of 10 evoke approximately linear increases in peak BOLD signals (Goodyear & Menon, 1998). Each condition combination was repeated twice per run in a random sequence, resulting in each subject experiencing a total of 18 visual trials per condition combination.

Each auditory trial consisted of a repeating sequence of two pure tones (440Hz and 550Hz, with a cycling rate of 10Hz) delivered over MR-compatible pneumatic earphones. Similar to the visual trials, there were a total of 9 possible trial conditions defined by the conjunction of three stimulus durations (100ms, 300ms, and 900ms) and three stimulus intensities (10% volume, 30% volume, and 90% volume). Each trial condition was repeated twice per run in a random sequence, resulting in each subject experiencing a total of 18

auditory trials per condition. To ensure that subjects remained attentive throughout the duration of each scan, subjects responded to each auditory stimulus by pressing a button on an MR-compatible button box (Current Designs, Inc.; Philadelphia, PA, USA) using their right index finger. Eye movements and closures were also monitored using an MR-compatible long-range Eyelink 1000 Plus eyetracker camera (SR Research; Ottawa, Ontario, Canada).

Two functional localiser runs (236s each) followed the main task. Contrast-reversing checkerboards (flicker rate: 10Hz) and oscillating pure tones (440Hz and 550Hz, with a cycling rate of 10Hz) were presented in alternating blocks of 16s separated by 12s fixation periods between blocks.

All but one participant completed the above session plan. This participant completed only six of the planned nine runs, and her data were excluded from further analysis. The final sample therefore included 15 participants.

MRI data acquisition and pre-processing

MRI data were acquired using a Siemens 3T MAGNETOM Prisma MRI scanner (Siemens, Erlangen, Germany) and a 32-channel head coil. Functional scan parameters were adapted from those used for the Human Connectome Project (HCP) and were chosen to ensure that full-brain coverage, including the cerebellum, was achieved for each participant.

A 3D high-resolution (1 x 1 x 1 mm) T1-weighted magnetisation-prepared, rapid acquisition gradient echo (MPRAGE) pulse sequence was used to obtain whole-brain anatomical images for each participant and normalize individual subject data to a standard space. 128 1-mm thick contiguous sagittal slices (0.5 mm skip; 1 x 1 mm in-plane resolution) were acquired with a repetition time (TR) of 2300 ms, an effective echo time (TE) of 2.22 ms, a flip angle (FA) of 8°, and a 260 x 260 mm field of view (FOV).

Functional MRI data were acquired using a multiband echoplanar imaging (MB-EPI; CMRR release R2015) sequence with a MB acceleration factor of 8 (Demetriou et al., 2018). 378 whole-brain images were obtained for each localiser run while 624 images were acquired for each task run. T2*-weighted images were acquired using a TR of 625 ms, a TE of 33.2 ms, and a FA of 50°. 64 interleaved slices (thickness = 2.5 mm, no gap) were collected using a 220 x 220 mm FOV (imaging matrix = 64 x 64). The voxel size was thus 2.5 x 2.5 x 2.5 mm^3.

Three task runs and one localizer run were discarded due to technical difficulties during data acquisition. The remaining task and localiser data were pre-processed using a previously published pre-processing pipeline (Li et al., 2019; Kong et al., 2019). The initial four images from each run were removed to aid with BOLD signal stabilisation. Motion correction using FSL's MCFLIRT was then applied such that runs with more than 50% of the frames removed were discarded to ameliorate any contributions of head motion. (This step excluded no runs from the present dataset.) FSL's bbregister function was used for intrasubject registration of the T1 anatomical images to the T2*-weighted images. The best run for each subject was used as the registration file across all functional runs. Subsequently, motion parameters and their derivatives were regressed out using global signal regression and censored frames were interpolated (Power et al., 2014). Linear trends were removed and the images were spatially smoothed using an isotropic Gaussian kernel of 6 mm (FWHM). Finally, the images were down sampled and projected to volumetric MNI 2-mm space. These volumetric outputs were subsequently submitted to further analysis in using FSL Release 6.0 (Jenkinson et al., 2012).

Region of interest (ROI) definitions

To account for between-subject differences in BOLD response locations, we sought to define regions-of-interest (ROIs) from each participant's functional localisers. Each localiser

run was first submitted on a per-subject basis to a first-level analysis in FEAT. Stimulation and fixation blocks were modelled using boxcar functions. Model regressors were created by convolving these boxcar functions with a single gamma function (SD = 3s, mean lag = 6s). The first-level localiser analyses were next combined in a fixed-effects higher-level analysis for each modality and participant, with voxel clustering set at $z > 3.1$ and $p < .05$.

Visual stimulation yielded robust activation clusters in the medial posterior occipital lobes of both the left and right hemispheres. Across participants, there was considerable variation in the spatial location of peak voxel activity. We created per-subject spherical ROI masks with a maximum radius of 6mm. Using FSLeyes, contiguous voxels surrounding the peak voxel were selected within the left and right primary visual cortex. These ROIs were constrained by the anatomical boundaries of primary visual cortex (BA 17 and 18) as defined by the Juelich Histological Atlas (Amunts et al., 2000) with a minimum probability threshold of 1%. The Juelich Histological Atlas is probabilistic; each ROI achieved at least 60% likelihood (left hemisphere, 82% in the right) of overlapping with primary visual cortex.

Functional localization was less successful for auditory localizer runs. Only 6 participants evidenced clusters in the lateral temporal lobes at the chosen threshold of $z > 3.1$ and $p < .05$. Consequently, auditory ROIs were constructed from atlas locations at the group level. ROI masks for the left and right primary auditory cortex were defined from the parcellation of sensory cortices in the Juelich atlas (Eickhoff et al., 2005). The resulting masks demonstrated a substantial degree of overlap with the few clusters of auditory cortex activity observed from the functional localizers. For comparison purposes, such atlas-based ROIs were also defined for primary visual cortex.

Event-related timecourses

Raw BOLD signal values were extracted from each subject's pre-processed 4-D functional data. These values were averaged across the voxels within each ROI. The resulting timecourse vectors were imported into MATLAB 2018b (Mathworks Inc., Natick MA) for temporal alignment and segmentation. Visual and auditory stimulus presentation times were extracted from scanner logs and converted to the volume of each event onset. Each region-specific BOLD signal vector was segmented into a 20 s event-related timecourse, starting 5 s prior to event onset. Each timecourse terminated 15 s after the event, thereby allowing the hemodynamic response to unfold without contamination from subsequent events or their anticipation.

Raw signal values were next converted into percent signal change values for each event-related time-course. Due to unexpectedly high cross-contamination (e.g. auditory responses in primary visual cortex), the baseline for each event was defined as the measured BOLD signal 1.25s after event onset. At 1.25s after event onset, preceding impulses would have fully dissipated, but responses to the present event were yet to unfold. Event-related time-courses were averaged across hemispheres because stimuli had been presented across visual hemifields or binaurally. Responses across hemispheres generally looked similar.

Expected responses from Gamma-based hemodynamic models

To generate expected hemodynamic responses, we created forward models for each experimental condition combination. Stimulus durations and intensities were modelled via boxcar functions, with intensity levels (1%, 10% and 100% contrasts; 10%, 30% and 90% volume) coded as linearly-increasing neural activation strengths (i.e., 1, 2, 3). These boxcar functions were then convolved with a canonical hemodynamic response function (Friston et al., 1995) to produce the predicted hemodynamic response curves. The canonical HRF is a double-Gamma function with six parameters (response delay, undershoot delay, response

dispersion, undershoot dispersion, ratio of response to undershoot, onset, and kernel length) as implemented in SPM (Friston et al., 1995). The starting values for each of these parameters followed the SPM defaults. Predictions from these models are presented alongside the relevant empirical results.

Fitting Gamma-based hemodynamic models to empirical data

Parameterised canonical hemodynamic response functions were also used for visualization and statistical analysis of the empirical data. Event-related average BOLD time-courses were first aggregated by condition combination. For visualization, each aggregated sample included all participants, which amounted to approximately 4800 discrete events (15 subjects x 9 runs x 2 ROIs x 9 conditions x 2 events per condition). For standard statistical analyses, aggregates were created for each participant separately. Bootstrap-based statistical analysis used aggregates created by sampling each participant's data with replacement; the motivation and procedures for this approach are detailed below.

For each aggregated data set, the best-fit canonical hemodynamic response function (double Gamma, adapted from the SPM implementation (Friston et al., 1995)) was found using a non-linear least squares algorithm implemented in MATLAB 2018b (Mathworks Inc., Natick MA). The response function had four free parameters: peak amplitude, response onset, peak delay, and peak dispersion. This approach was similar to Wager et al. (2005)'s, save the addition of peak dispersion as a free parameter. Its inclusion produced modelled hemodynamic response curves that were better aligned with the event-related average time-courses, particularly in the response's falling phase.

Obtained parameter estimates were used to generate hemodynamic response curves. For statistical analyses, parameters were then derived from these curves. Peak magnitude and peak latency were defined as the maximum BOLD percent signal change and its

corresponding timepoint. Response onset latency was defined as the time at which the hemodynamic response reached 10% of the peak amplitude. Peak dispersion was defined as the full-width half-max (FWHM) value of the curve. Group means, as well as within-subject standard deviations and 95% confidence intervals (CIs), were subsequently calculated for all four measures (Cousineau, 2005).

Bootstrap resampling and statistical analyses

Conventional inferential statistics rely on metric estimates from individual subject data, with each subject's measures assumed to be equally reliable (Suchow et al., 2013). Neuroimaging measures, however, exhibit considerable variance across timepoints and participants, rendering it challenging to derive good estimates of hemodynamic response onset and peak latency—even after using the Gamma-based fitting procedures described above. A potential solution comes from EEG analyses, in which evoked response potential timing is also important but difficult to estimate (Liesefeld, 2018). Group-averaged timecourses increase the signal-to-noise ratio, thereby allowing for the estimation of reliable sample parameters (Asplund et al., 2010). Variance of the estimates is obtained through resampling.

Here we used bootstrap resampling to estimate parameter means and variances, in addition to calculating traditional by-participant estimates. For our bootstrap procedure, 10,000 "pseudo-samples" of equivalent size to the original dataset were constructed by sampling by-participant data (mean event-related timecourses for each condition combination) at random and with replacement (Efron & Gong, 1983). Each bootstrap sample was created separately for each condition combination. For each bootstrap sample, the best-fit canonical hemodynamic response function was found using a non-linear least squares

algorithm, which was then used to generate hemodynamic response curves from which the bootstrapped parameter values were obtained.

Statistical analysis of bootstrapped and non-bootstrapped parameter values for the fitted gamma-based hemodynamic responses followed a largely similar workflow to facilitate comparisons of results by condition and usage of resampling method. For each unique parameter, means, standard deviations (SD), standard errors (SE) and 95% confidence intervals (CI) were derived across all nine study condition combinations for each stimulus modality. Statistics of pairwise differences (means, standard deviations and 95% confidence intervals) were computed between stimulus durations at each intensity level and between stimulus intensities at each duration level.

Unless stated otherwise, figures and reported findings in Experiment 1 reflect bootstrapped parameter estimates. Standard parameter estimates were used for calculating effect sizes (e.g., Cohen's *d*). Figures and reported findings for Experiment 2 reflect conventional by-participant estimates.

Results

Visual stimulus duration and intensity effects were consistent with predicted HRFs

The simulated BOLD responses illustrated the expected effects of stimulus intensity and duration. When stimulus duration increased but intensity was held constant, both peak magnitude and peak latency increased (Figure 1, top row). The empirical results were broadly consistent with these simulated effects. At maximal stimulus contrast (100%), increases in visual stimulus duration led to increases in both peak magnitude and peak latency of the fitted hemodynamic responses (Figure 1, bottom row). Each increase in stimulus duration led to a statistically significant increase in average peak magnitude, which doubled across the tested stimulus duration range (large effect size, d = 1.44). Pairwise changes in peak latency,

however, were too variable to reach statistical significance. The substantial 0.59 s increase (95% CI [-0.09 1.33], d = 0.73) in peak latency across the tested range accorded with predictions, but its variability was likewise large.

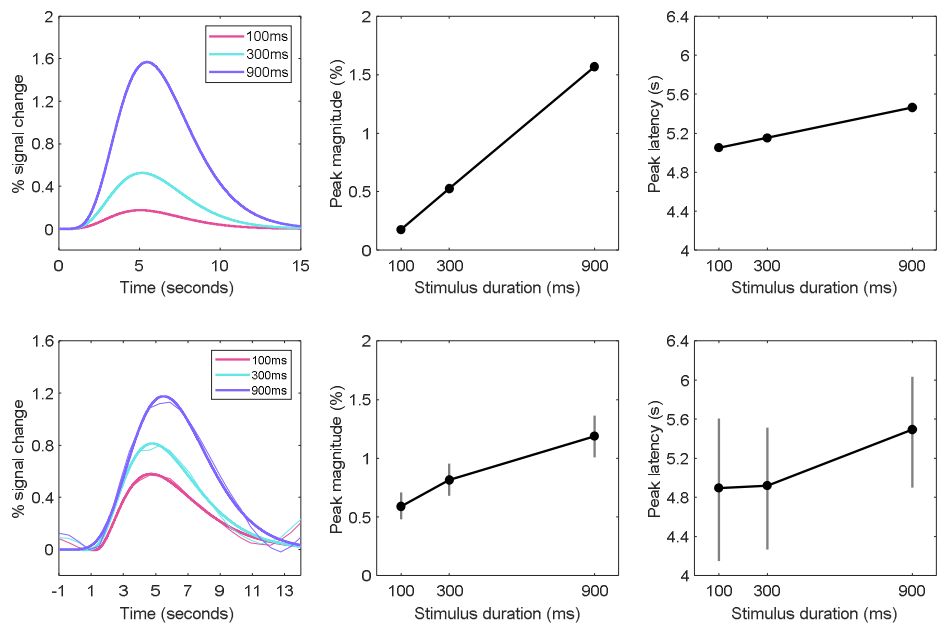

Figure 1. Predicted (top row) and observed (bottom row) effects of stimulus duration on the overall hemodynamic response (left column), peak magnitude (middle column), and peak latency (right column) evoked by maximum-contrast (100%) stimuli in primary visual cortex. Observed hemodynamic responses were obtained by averaging across participants. Observed peak magnitude and peak latency estimates and 95% CIs were derived from bootstrap samples. Increases in stimulus duration led to increases in both the peak magnitude and peak latency of the event-related hemodynamic response, albeit with high variance in the latter metric.

The simulated BOLD responses also illustrated the expected contrast between manipulations of stimulus duration and intensity. When stimulus intensity increased but duration was held constant, only the magnitude of the hemodynamic response increased

(Figure 2, top row). The empirical effects were again broadly consistent with the simulations. At the 900 ms stimulus duration, peak magnitude increased as stimulus contrast increased, doubling across the tested range (large effect size, d = 1.29; Figure 2, bottom row). Peak latency showed only small (0.20 s, 95% CI [-0.49 1.00], d = -0.02) and non-significant changes across the tested range.

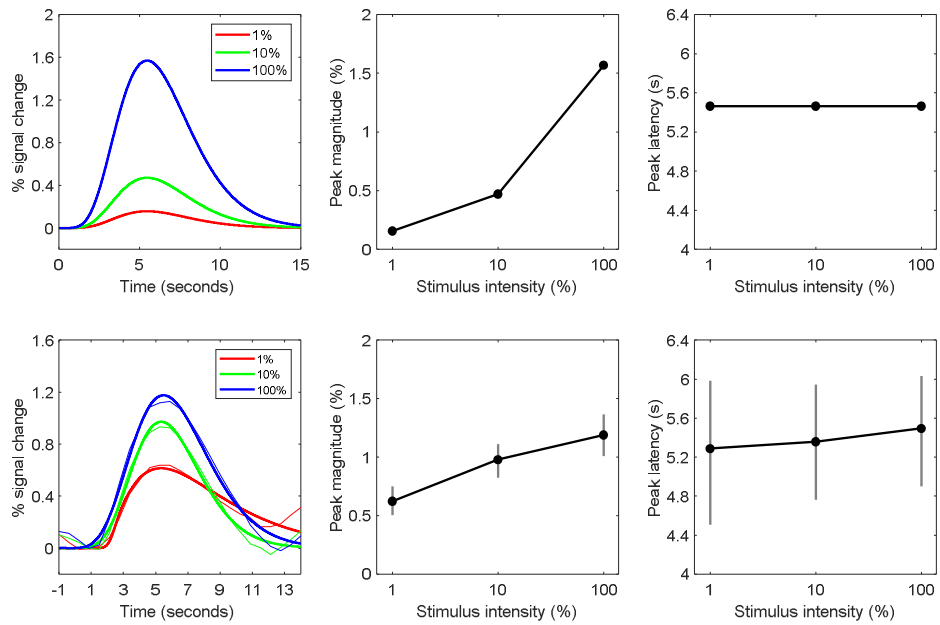

Figure 2. Predicted (top row) and observed (bottom row) effects of stimulus intensity on the overall hemodynamic response (left column), peak magnitude (middle column), and peak latency (right column) evoked by 900 ms stimuli in primary visual cortex. Increasing stimulus intensity (i.e., contrast) led to increases in peak magnitude, but not peak latency, of the event-related hemodynamic response.

Notably, the standard analytical approach produced a mean peak latency of 6.08 s (SD = 1.70, 95% CI [5.22, 6.94]) for the 1% contrast condition, as compared to 5.45 s (SD = 0.79, 95% CI [5.05, 5.85]) for the 10% contrast condition. Inspection of the individual fits revealed

peak latencies in excess of 6 s for 5 of 15 participants in the 1% contrast condition, for which overall activation was low. This situation illustrates the difficulty of deriving robust temporal parameters in the presence of noise that overwhelms the signal (Liesefeld, 2018). The bootstrap approach is a potential remedy to this problem. Indeed, the bootstrapped peak latency estimates for 900 ms stimuli differed by just 0.2 s across the three levels of stimulus contrast, perhaps reflecting greater metric stability. Regardless, metrics derived from either approach were not consistent with increasing peak latency with increasing stimulus contrast.

Latency estimates were highly variable at lower visual stimulus intensities

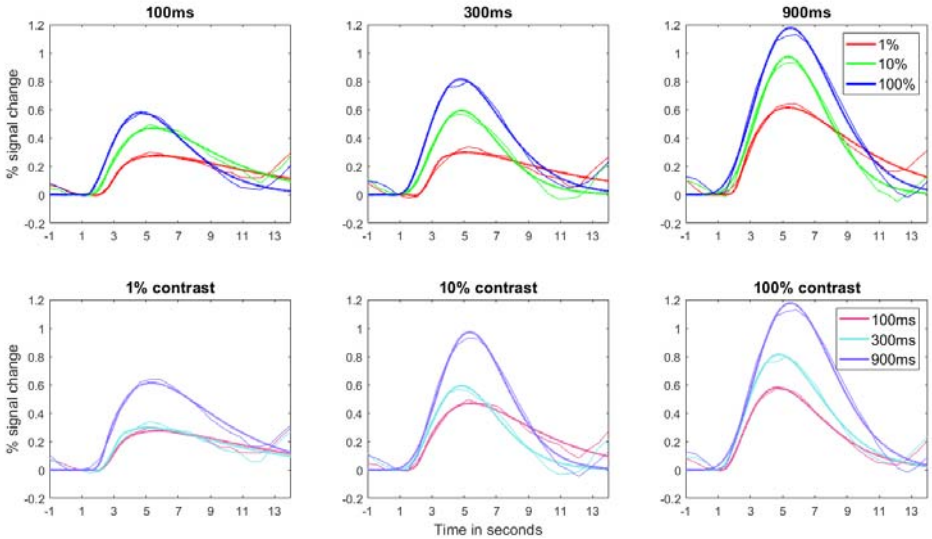

Figure 3. Event-related hemodynamic responses in primary visual cortex across increasing levels of stimulus intensity at each level of stimulus duration (top row) and increasing levels of stimulus duration at each level of stimulus intensity (bottom row).

Peak magnitude was sensitive to manipulations of stimulus duration at lower contrast levels (Figure 3, bottom row; Figure 4). Similarly, peak magnitude reflected contrast manipulations with shorter-duration stimuli (Figure 3, top row; Figure 4). For example, the

difference between 1% and 100% contrast for 300 ms stimuli had a large effect size (d = 1.31), comparable to the results for 900 ms stimuli.

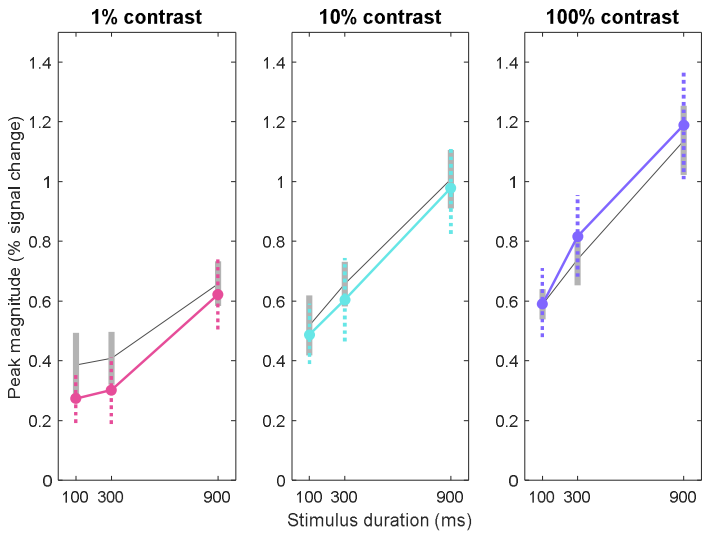

Figure 4. Peak magnitude estimates across visual stimulus durations and contrasts. Increases in stimulus durations were associated with increases in peak magnitudes of the corresponding event-related hemodynamic responses at each level of visual stimulus contrast. Bootstrapped estimates and error bars are displayed in colour; non-bootstrapped estimates and error bars are plotted in grey. Error bars represent 95% CIs. Significant increases in mean peak magnitude can be observed between responses to 300ms and 900ms stimuli at all levels of stimulus contrast.

At lower levels of stimulus contrast, peak latency differences—or the lack of them—could not be established (Figure 3, Figure 5). In particular, the observed responses in the 1% contrast condition often had very low amplitudes or deviated from expected HRF shapes. Accordingly, the variance for these estimates was high. Bootstrap resampling did not significantly ameliorate this problem.

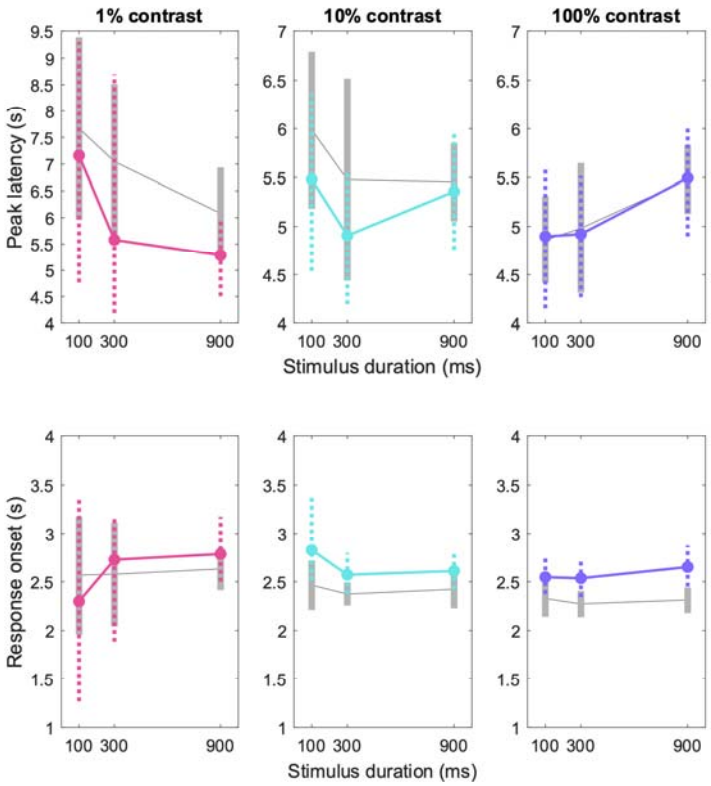

Figure 5. Peak latency and response onset estimates across visual stimulus durations and contrasts. Bootstrapped estimates and error bars are displayed in colour; non-bootstrapped estimates and error bars are plotted in grey. Error bars represent 95% confidence intervals of the mean. There were negligible effects of stimulus duration on the response onsets (bottom row) of the corresponding event-related hemodynamic responses at each level of stimulus contrast.

Auditory stimulus duration and intensity effects were consistent with predicted HRFs

The empirical results were broadly consistent with the forward model predictions (Figure 6, top row), albeit less convincingly so than in the visual domain. When stimulus volume was held constant at 90% of the maximum, increases in auditory stimulus duration led to increases in both peak magnitude and peak latency of the fitted hemodynamic responses (Figure 6, bottom row). Across the tested range (100 to 900 ms), peak magnitude increased by 0.19% (SE = 0.09, 95% CI [0.03 0.36], d = 1.04) and peak latency increased by 0.27 s (SE = 0.20, 95% CI [-0.13 0.65], d = 0.55).

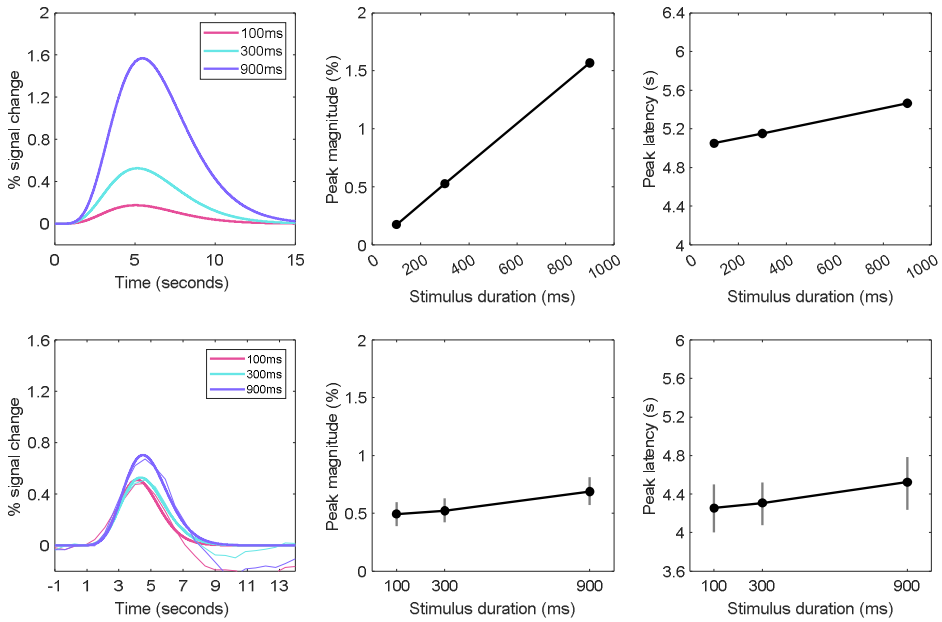

Figure 6. Predicted (top row) and observed (bottom row) effects of stimulus duration on the overall hemodynamic response (left column), peak magnitude (middle column) and peak latency (right column) evoked by stimuli at the maximum tested volume (90%) in primary auditory cortex. Observed peak magnitude and peak latency estimates and 95% CIs were derived from bootstrap samples.

Conversely, when auditory stimulus intensity increased but duration was held constant at 900 ms, peak magnitude moderately increased while peak latency showed no evidence of an increase. Indeed, across the tested range, peak magnitude increased by 0.23% (SE = 0.08, 95% CI [0.07 0.39], d = 0.99), but peak latency (non-significantly) decreased by 0.25s (SE = 0.21, 95% CI [-0.67 0.16], d = -0.17).

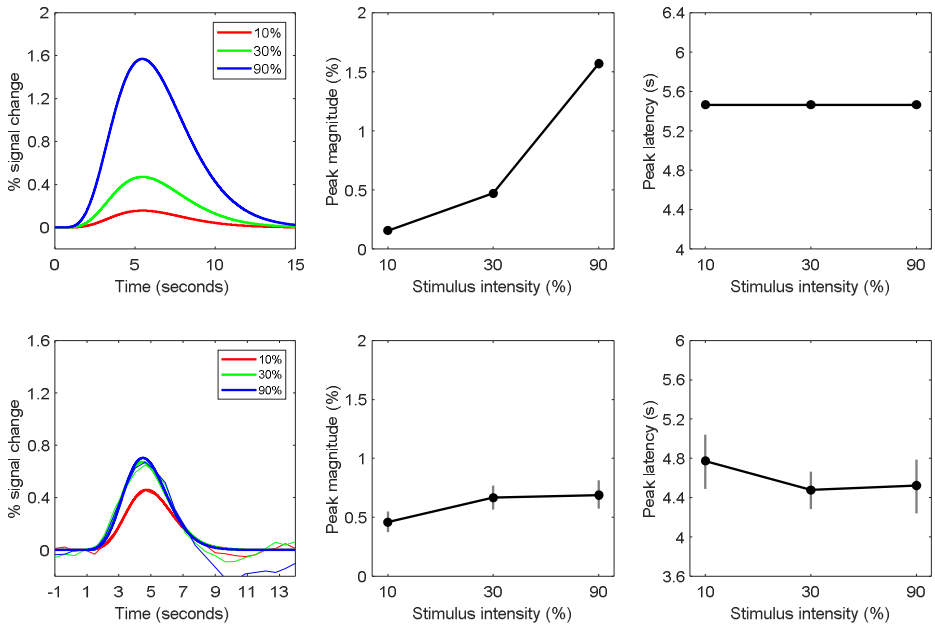

Figure 7. Predicted (top row) and observed (bottom row) effects of stimulus intensity on the overall hemodynamic response (left column), peak magnitude (middle column) and peak latency (right column) evoked by 900ms stimuli at primary auditory cortex. Observed peak magnitude and peak latency estimates and 95% CIs were derived from bootstrap samples. Increasing stimulus intensity (i.e., volume) leads to increases in the peak magnitude, but not the peak latency, of the event-related hemodynamic response.

Latency changes were small at lower auditory stimulus intensities

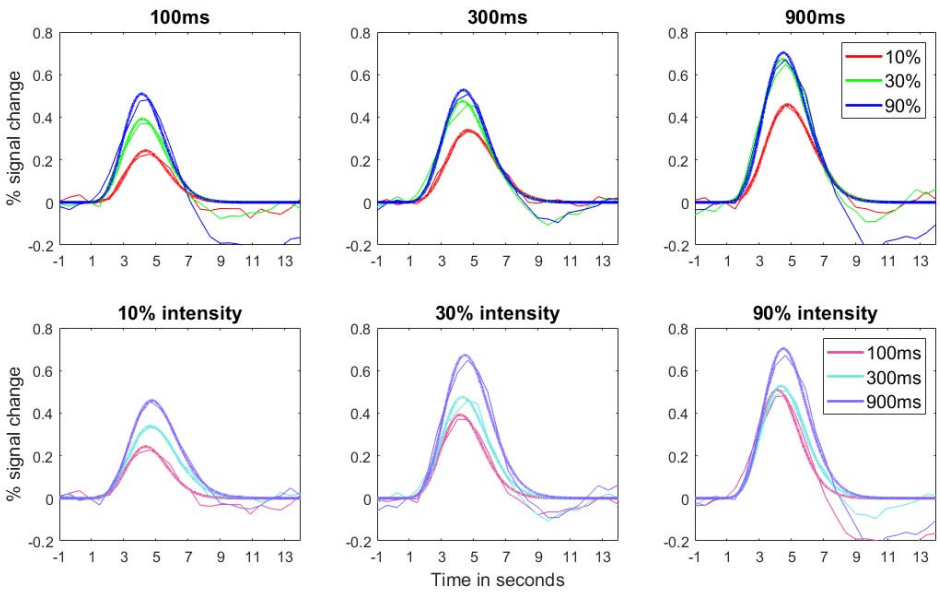

Figure 8. Event-related hemodynamic responses in primary auditory cortex across increasing levels of stimulus intensity at each level of stimulus duration (top row) and increasing levels of stimulus duration at each level of stimulus intensity (bottom row).

At lower stimulus intensities, peak magnitude was sensitive to manipulations of auditory stimulus duration (Figure 8, bottom row; Figure 9). Similarly, peak magnitude reflected intensity (volume) manipulations even with short-duration stimuli (Figure 8, top row; Figure 9). Some peak magnitude differences were less pronounced than their visual counterparts. That difference could reflect the stimulus intensity factors used (10x for visual, 3x for auditory) or the influence of background acoustic noise on baseline auditory cortex activation as well as perception.

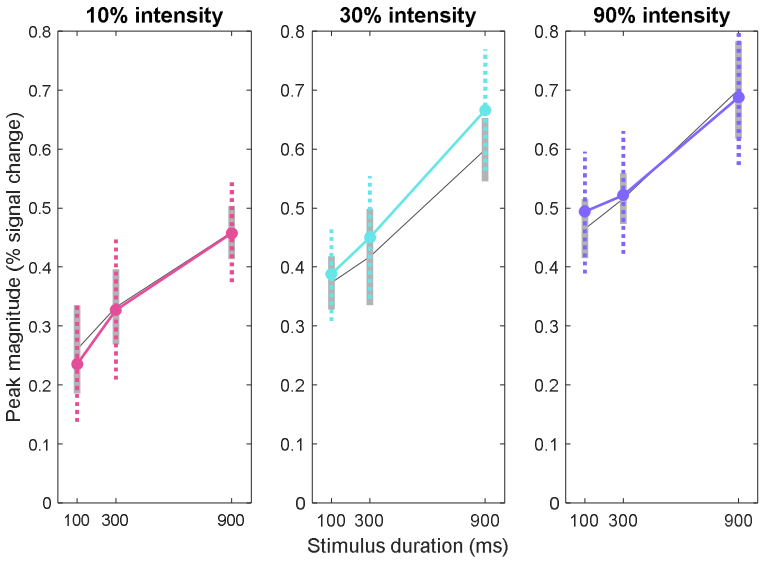

Figure 9. Peak magnitude estimates across auditory stimulus durations and contrasts. Increases in stimulus duration were associated with suggestive but non-significant increases in peak magnitudes of the corresponding event-related hemodynamic responses at each level of visual stimulus contrast. Bootstrapped estimates and error bars are displayed in colour; non-bootstrapped estimates and error bars are plotted in grey. Error bars represent 95% confidence intervals of the mean.

Unlike in the visual domain, however, peak latency estimates remained relatively robust even at low stimulus intensity levels (Figure 10). Few individuals showed very low activation, even in the low volume conditions. The bootstrapping procedure appeared to be more helpful here as well, perhaps because its use ameliorated random noise contributions instead of apparent deviations from standard HRFs.

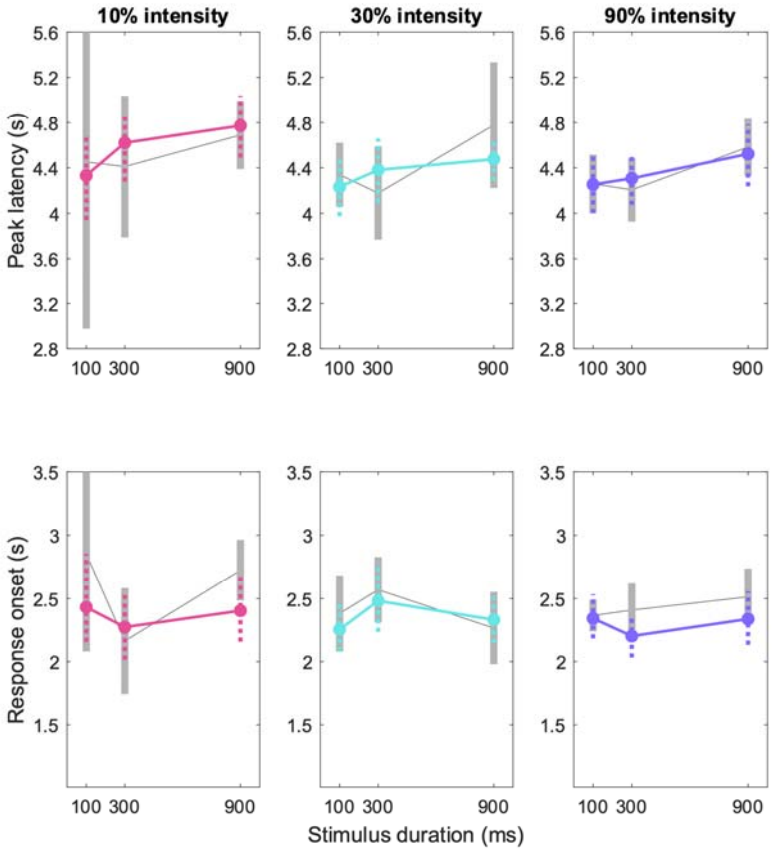

Figure 10. Peak latency and response onset estimates across auditory stimulus durations and contrasts. Increases in stimulus duration led to non-significant but positive trends in mean peak latencies of event-related hemodynamic responses at primary auditory cortex across all levels of stimulus intensity. Bootstrapped estimates and error bars are displayed in colour; non-bootstrapped estimates and error bars are plotted in grey. Error bars represent 95% confidence intervals of the mean. There were negligible effects of stimulus duration on response onset (bottom row) at each level of stimulus intensity.

Experiment 1 Discussion

Experiment 1 investigated whether the BOLD signal's peak magnitude and peak latency were sensitive to manipulations of stimulus duration and intensity. These parameters were estimated from double-Gamma hemodynamic functions fitted to individual subject data or bootstrap samples. For both auditory and visual stimuli, the results were largely consistent with predictions and previous empirical findings. Increases in stimulus duration or intensity

produced non-linear increases in peak magnitude (Liu & Gao, 2000; Vazquez & Noll, 1998; Wager et al., 2005; Yeşilyurt et al., 2008). In contrast, only increases in stimulus durations appeared to affect peak latency.

For most comparisons of interest, however, high parameter variance (wide 95% confidence intervals) precluded drawing strong inferences about the effects of experimental manipulations on those parameters. In particular, peak latency differences could not be confidently established for stimulus duration changes. Moreover, there remained significant variability in the estimated hemodynamic response parameters at lower stimulus intensities, even when using bootstrap resampling approaches. This result may reflect a limit in the extent to which the canonical HRF can characterise the BOLD responses evoked by the brief periods of visual and auditory stimulation. When analysing such responses, canonical HRF models may demonstrate reduced sensitivity to activation, as rapid signal fluctuations and temporal heterogeneity across voxels lead to responses that are both smaller in amplitude and more variable across trials (Polimeni & Lewis, 2021).

Experiment 2

In Experiment 1, brief visual and auditory events often produced hemodynamic responses with low signal-to-noise ratios, thereby leading to parameter estimates with a high degree of variability. To address this limitation, a follow-up study was conducted with several modifications to increase the reliability of these parameter estimates. Foremost, stimulus duration was increased to ensure robust hemodynamic responses; the small duration differences between stimuli (200 or 600 ms) were preserved so that small relative differences could be investigated. The number of trials per condition was also increased (and number of conditions reduced) as was the TR, with both changes intended to increase signal-to-noise of the collected data. Aside from these changes, Experiment 2 retained many of the features of

Experiment 1, utilising the same types of visual and auditory stimuli, retaining the same degree of difference between each level of stimulus duration, and administering these stimuli at similar intervals within a similar run length. We predicted that duration and intensity changes would both affect peak magnitudes, whereas only duration changes would affect peak latency.

Materials and Methods

Experimental procedures

The procedures were identical to those in Experiment 1 except as noted. Fourteen members from the NUS community (5 female, 13 right-handed, mean (±SD) age = 27.5±2.36 years) were recruited for the study. They were paid 30 SGD for their participation.

Participants underwent a 60-minute slow event-related main task, conducted across 8 runs of 350 s each. Each run consisted of 16 visual trials and 16 auditory trials. The same contrast-reversing and alternating-tone stimuli from Experiment 1 were used.

There were four visual trial conditions. For three of them, the visual stimuli were presented at maximal intensity (100% contrast) for differing stimulus durations (1000ms, 1200ms and 1800ms). In the fourth condition, the visual stimuli were presented at a relatively lower intensity (10% contrast) for 1200ms. Each trial condition was repeated four times per run in a random sequence, resulting in each subject experiencing a total of 32 visual trials per condition over the course of the entire task.

As with the visual trials, there were four auditory trial conditions. In three of these conditions, the auditory stimuli were presented at a high intensity (90% volume) for differing stimulus durations (1000ms, 1200ms and 1800ms). In the fourth condition, auditory stimuli were presented at a relatively lower intensity (30% volume) for 1200ms. Each trial condition was repeated four times per run in a random sequence, resulting in each subject experiencing

a total of 32 auditory trials per condition over the course of the entire task. To ensure that subjects remained attentive throughout the duration of each scan, subjects responded to each auditory stimulus by pressing a button on an MR-compatible button box (Current Designs, Inc.; Philadelphia, PA, USA) using their right index finger.

Two functional localiser runs (236 s each) followed the main task. Contrast-reversing checkerboards (flicker rate: 10Hz) and oscillating pure tones (440Hz and 550Hz, with a cycling rate of 10Hz) were presented in alternating blocks of 16s separated by 12s fixation periods between blocks.

All but two participants completed the above session plan. As these participants did not complete at least 80% of the planned runs, their data was excluded from further analysis. The final sample therefore consisted of 12 participants.

MRI data acquisition, pre-processing, and ROI definition

The MRI data acquisition approach was the same as in Experiment 1 with the following exceptions. Functional MRI data were acquired using a multiband echoplanar imaging (MB-EPI; CMRR release R2015) sequence with a MB acceleration factor of 6. 236 whole-brain images were obtained for each localiser run while 350 images were acquired for each task run. T2*-weighted images were acquired using a TR of 1000 ms, a TE of 31.8 ms, and a FA of 50°. 60 interleaved slices (thickness = 2.5 mm, no gap) were collected using a 220 x 220 mm FOV (imaging matrix = 64 x 64). The voxel size was thus 2.5 x 2.5 x 2.5 mm^3.

All task and localiser data were pre-processed using the same procedures as in Experiment 1. None of the runs were excluded from the dataset during pre-processing. Visual stimulation from the localisers yielded robust activation clusters in the medial posterior occipital lobes of both the left and right hemispheres. Spherical ROI masks were created for each subject with a maximum radius of 6mm. Using FSLeyes, continguous voxels

surrounding the peak voxel were selected within the left and right primary visual cortex. These ROIs were constrained by the anatomical boundaries of primary visual cortex (BA 17 and 18) as defined by the Juelich Histological Atlas (Amunts, Malikovic, Mohlberg, Schormann, & Zilles, 2000).

Auditory stimulation from the localisers also yielded robust activation clusters at the lateral temporal lobes of both the left and right hemispheres in all but two participants. Spherical ROI masks (maximum radius of 6 mm) were created for each subject who had robust auditory activation. For the remaining two participants, auditory ROIs were constructed from atlas locations.

Event-related timecourse extraction, Gamma-based hemodynamic models

In Experiment 2, the baseline for each event was defined as the BOLD signal level occurring 1s before event onset. For visualization purposes, event-related time-courses were averaged by modality, collapsing across hemispheres because stimuli had been presented across visual hemifields or binaurally. As in Experiment 1, responses across hemispheres looked similar. Gamma-based hemodynamic response fitting, parameter extraction, and analysis procedures were similar to those from Experiment 1.

Results

Peak latency differentiates visual stimulus duration and intensity manipulations

As predicted, increases in visual stimulus durations led to significant increases in both peak magnitudes and peak latencies of the hemodynamic response (Figure 11, top left; Figure 12). Peak magnitude increased by 0.11 (SD = 0.18, 95% CI [0.01 0.22]) as stimulus duration increased by 200 ms (from 1000 to 1200 ms), and it increased by 0.38 (SD = 0.20, 95% CI [0.26 0.49]) as stimulus duration increased by 600 ms (from 1200 to 1800 ms). The overall

peak magnitude increase of 0.50 (SD = 0.16, 95% CI [0.40 0.59]) was similar to the increase observed in Experiment 1 for the same stimulus duration change.

Critically, increases in visual stimulus durations also led to significant increases in peak latency. Peak latency increased by 0.18 s (SD = 0.26, 95% CI [0.03 0.33]) as stimulus duration increased by 200 ms (from 1000 to 1200 ms), and it increased by 0.33 s (SD = 0.49, 95% CI [0.04 0.60]) as stimulus duration increased by 600 ms (from 1200 to 1800 ms). The overall peak latency increase of 0.50 s (SD = 0.18, 95% CI [0.40 0.61]) was similar to the increase observed in Experiment 1; the key difference was the much lower variance in Experiment 2.

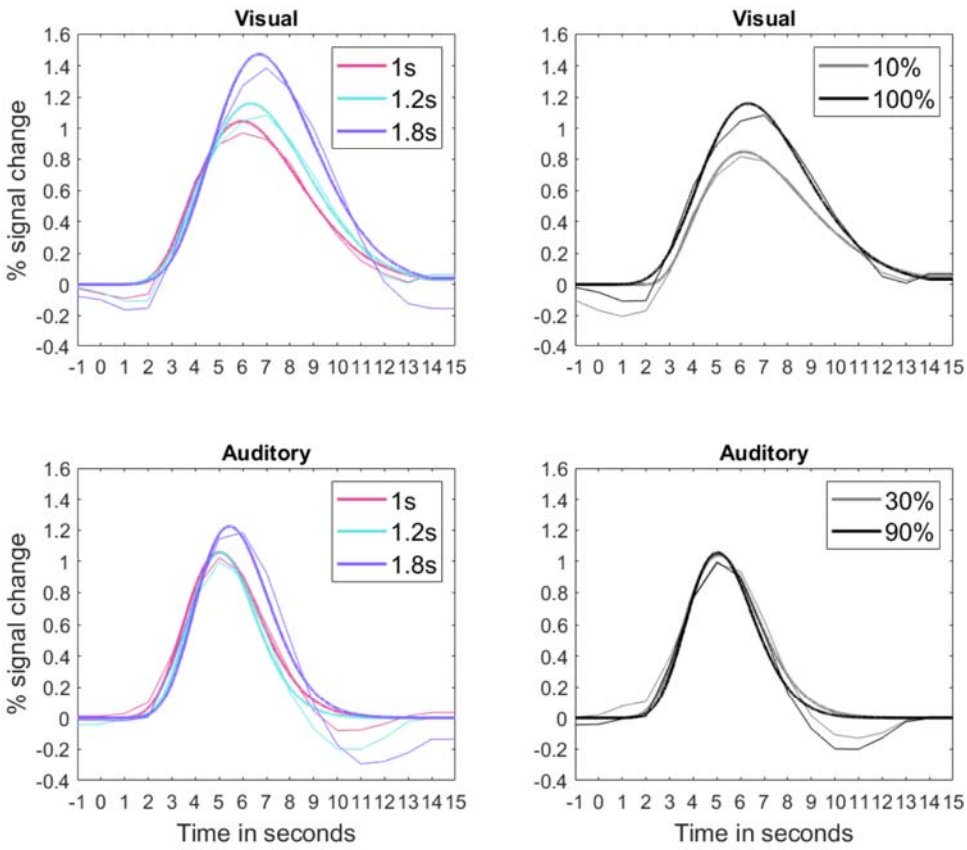

Figure 11. Fitted event-related hemodynamic responses in primary visual cortex (top row) and primary auditory cortex (bottom row). Effects of visual and auditory stimulus durations

of 1000 ms, 1200 ms and 1800 ms at maximal stimulus intensity (left column) and across two levels of visual and auditory stimulus intensity at the 1200 ms stimulus duration (right column). Note the peak latency changes for duration manipulations but not for different intensities.

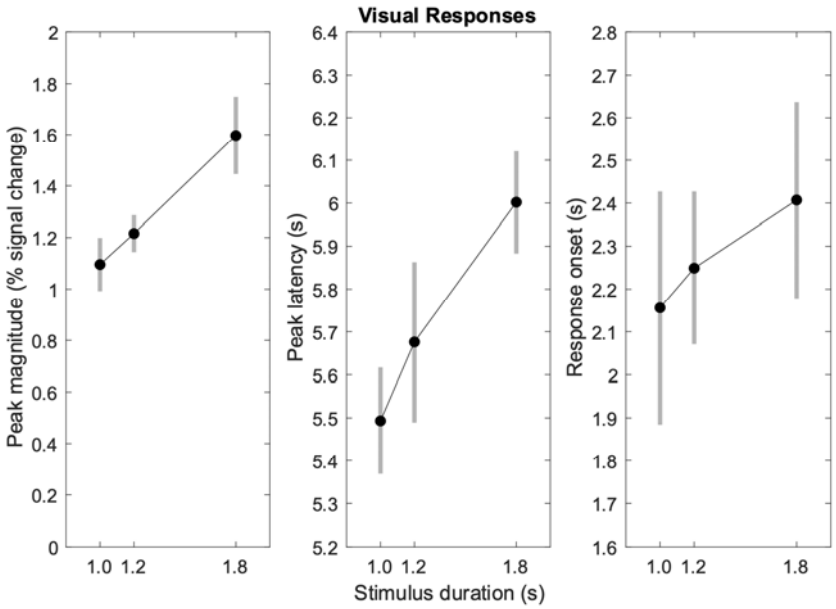

Figure 12. Estimated HRF parameters from visual stimulation across different durations. 600ms (1200ms-1800ms) and 800ms (1000ms-1800ms) increases in visual stimulus durations led to significant increases in the mean peak magnitude (left column) and peak latency (middle column) of their corresponding fitted hemodynamic responses. There were negligible effects of stimulus duration on response onset (right column). Error bars represent 95% confidence intervals of the mean.

In contrast, increases in stimulus intensity affected only the HRF's peak magnitude, not its peak latency (Figure 11, top right; Figure 13). The increase in stimulus contrast from 10% to 100% produced an 0.18% (SD = 0.29, 95% CI [0.02 0.34]) increase in peak magnitude. This same stimulus change, however, produced a non-significant change of -0.20s (SD = 0.49, 95% CI [-0.48 0.07]) in peak latency.

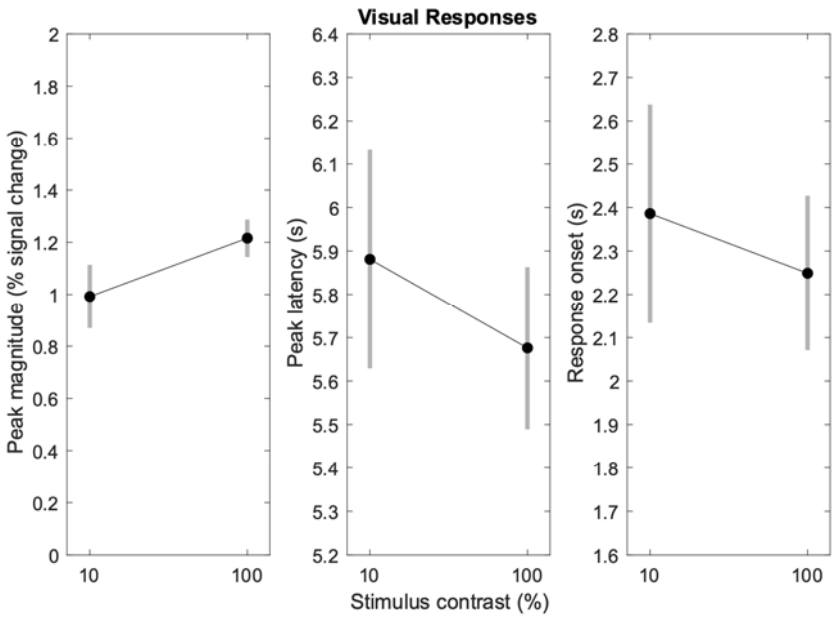

Figure 13. Estimated HRF parameters from visual stimulation across different intensities. Increasing visual stimulus intensity (contrast) led to a significant increase in the mean peak magnitude (left column) of the corresponding fitted hemodynamic responses. There were negligible effects of stimulus duration on peak latency (middle column) and response onset (right column). Error bars represent 95% confidence intervals of the mean.

Peak latency differentiates auditory stimulus duration and intensity manipulations

Increases in auditory stimulus durations led to significant increases in both peak magnitudes and peak latencies of the hemodynamic response (Figure 11, bottom left; Figure 14). Peak magnitude increased by 0.10 (SD = 0.16, 95% CI [0.003 0.19]) as stimulus duration increased by 200 ms (from 1000 to 1200 ms), and it increased by 0.18 (SD = 0.24, 95% CI [0.04 0.31]) as stimulus duration increased by 600 ms (from 1200 to 1800 ms). The overall peak magnitude increase of 0.28 (SD = 0.09, 95% CI [0.22 0.33]) was similar to the increase observed in Experiment 1 for the same stimulus duration change, and both were smaller than their visual counterparts.

Critically, increases in auditory stimulus durations also led to largely significant increases in peak latency. Peak latency increased by 0.04 s (SD = 0.33, 95% CI [-0.15 0.22], so non-significant) as stimulus duration increased by 200 ms (from 1000 to 1200 ms), and it

increased by 0.55 s, (SD = 0.48, 95% CI [0.27 0.82]) as stimulus duration increased by 600 ms (from 1200 to 1800 ms). The overall peak latency increase of 0.59s (SD = 0.15, 95% CI [0.49 0.67]) was considerably larger and less variable than the increase observed in Experiment 1. It was comparable, however, to the peak latency increases observed for visual stimuli.

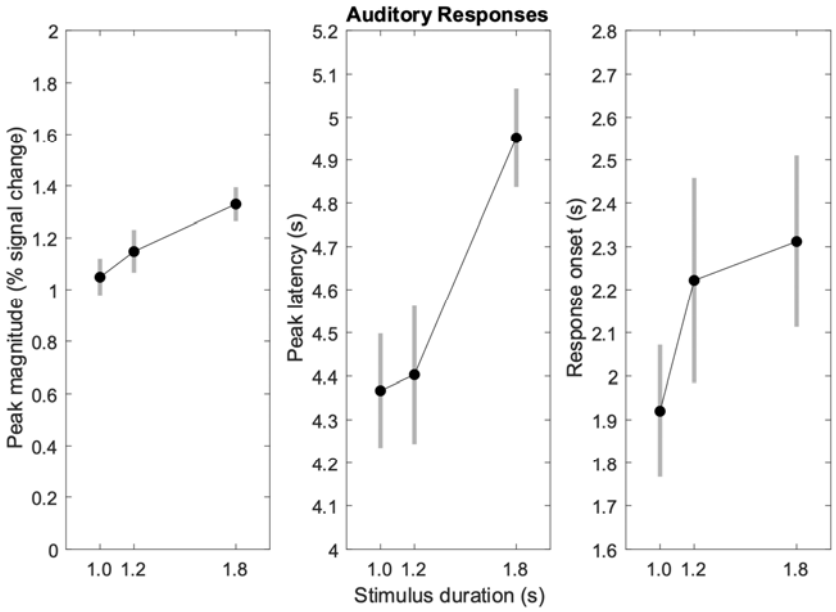

Figure 14. Estimated HRF parameters from auditory stimulation across different durations. 600ms (1200ms-1800ms) and 800ms (1000ms-1800ms) increases in auditory stimulus durations led to significant increases in the mean peak magnitude (left column) and peak latency (middle column) of their corresponding fitted hemodynamic responses. There were negligible effects of stimulus duration on response onset (right column). Error bars represent 95% confidence intervals of the mean.

Increases in stimulus intensity had no significant effects on either the HRF's peak magnitude or its peak latency (Figure 11, bottom right; Figure 15). The increase in stimulus volume from 30% to 90% of the speaker's maximum produced an 0.07% (SD = 0.18, 95% CI [-0.02 0.18]) increase in peak magnitude. This same stimulus change also produced a change of -0.11s (SD = 0.26, 95% CI [-0.26 0.04]) in peak latency.

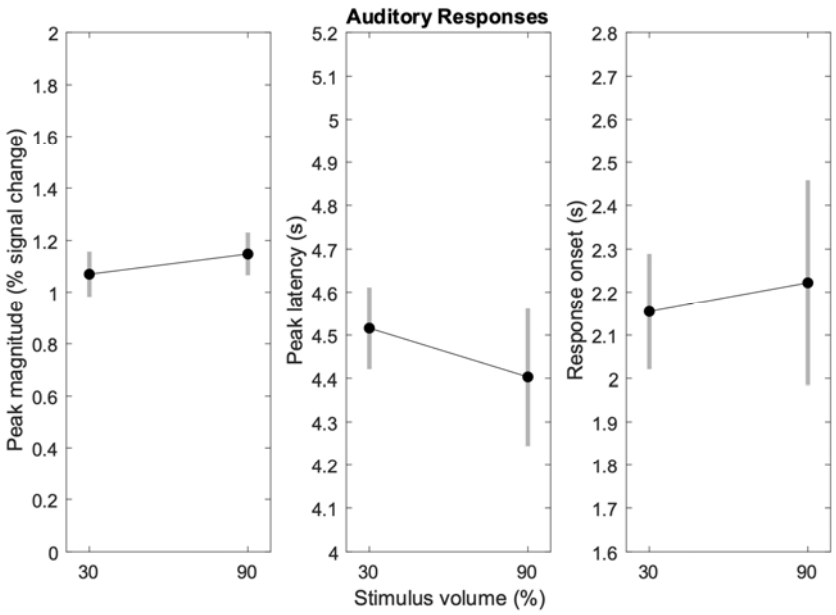

Figure 15. Estimated HRF parameters from auditory stimulation across different intensities. Increasing auditory stimulus intensity (contrast) led to negligible effects on the peak magnitude (left column), peak latency (middle column) and response onset (right column) of the corresponding fitted hemodynamic responses. Error bars represent 95% confidence intervals of the mean.

General Discussion

Summary of Findings

Differences in stimulus presentation durations affect both the peak magnitude and the peak latency of their corresponding event-related hemodynamic responses. In contrast, differences in the intensity of the stimuli lead only to differences in peak magnitude, not peak latency. The reported findings in Experiments 1 and 2 support these predictions and demonstrate that local event-related BOLD responses can precisely index sub-second differences in stimulus duration, perhaps to as low as 200 ms, when the underlying evoked signal is sufficiently robust.

Experiments 1 and 2 sought to explore whether visual and auditory responses associated with brief stimulus durations could be meaningfully distinguished from one another through hemodynamic response function parameters such as peak magnitude and peak latency. In Experiment 1, significant differences in average peak magnitude were observed at maximal stimulus intensity when comparing visual and auditory responses across stimulus durations. While these results suggested that even a 200 ms difference in stimulus duration could lead to a significant increase in mean peak magnitude of the event-related hemodynamic response across both perceptual domains, issues of high signal variability and relatively low degrees of signal change precluded our drawing strong timing-related conclusions. Experiment 2 addressed these issues through a series of changes to task-related and scanning parameters, producing robust and reliable responses to visual and auditory stimuli at durations of 1000, 1200, and 1800 ms. Even though the size of the duration increments in Experiment 2 were identical to those in Experiment 1 (200 and 600 ms), differences in both average peak magnitude and average peak latency were observed when comparing visual and auditory responses evoked by 1200 and 1800 ms stimuli, with some significant differences even for 200 ms duration changes.

Observed responses to short-duration stimuli support existing hemodynamic models

The present work extends the literature in several respects. Firstly, our findings not only corroborate the assertion that the effect of changes in stimulus intensity on event-related BOLD signal change can be dissociated from the effect of changes in stimulus duration (Grinband et al., 2008), but also build on previous work (Boynton et al., 2012) by extending the validity of such assertions to brief periods of stimulation of less than 2000 ms in duration and timing differences as low as 200 ms. Specifically, changes in stimulus intensity induce a non-linear scaling of peak magnitude without a concomitant increase in peak latency, whereas changes in stimulus duration affect peak magnitude and peak latency simultaneously.

Across both Experiments 1 and 2, we observed that the mean peak latencies associated with visual and auditory evoked responses did not differ significantly as a result of changes in stimulus intensity. We have provided further empirical evidence that peak amplitude and peak latency measures can be highly consistent across trials and subjects (Miezin et al., 2000), and that estimates of hemodynamic timing differences can be detected within intervals similar to those implied by cognitive events.

Our findings also highlight the relevance and utility of variable-epoch models in contrast-based fMRI analyses (Grinband et al., 2008). While variable-impulse models assume that the only aspect of the event-related hemodynamic response that meaningfully changes is the height of the response, variable-epoch models capture both the peak magnitude and the peak latency of the event-related BOLD signal (Lindquist et al., 2009). Though others have examined the properties of the hemodynamic response at similarly brief stimulus durations (Yeşilyurt et al., 2008), these studies have primarily focused on demonstrating how extremely small changes in stimulus duration affect the height of the evoked response, with relatively less attention paid to the peak latency parameter. Studies involving short-duration stimuli and sub-second timing differences have also at times failed to uncover significant differences in peak latency as a function of stimulus duration (Lewis et al., 2018). This study demonstrates that while conventional HRF models may indeed be inadequate at differentiating between evoked responses that involve sub-second stimulus presentations (Polimeni & Lewis, 2021), significant effects of relatively minute, sub-second differences in stimulus duration on the mean peak latency of the evoked hemodynamic response can be recovered by estimating hemodynamic response function parameters from the raw event-related BOLD signal using a non-linear least-squares optimisation algorithm. We contend that this approach is fairly straightforward, efficient, parsimonious, and easily interpretable, as the results of our analyses are ultimately determined by the degree to which measured task-evoked BOLD

activation fits an *a priori* hemodynamic response function with no further abstraction required. We suggest that temporal inferences can indeed be made from fMRI data using time-resolved methods of fMRI analysis such as mental chronometry (Tombu et al., 2011) and dynamic causal modelling (Stephan et al., 2007) with sufficient confidence. Although the BOLD response unfolds relatively slowly relative to its underlying neural causes, the fact that explicit manipulations of perceptual attributes do indeed lead to detectable hemodynamic differences implies that these effects would also be applicable to other cognitive operations and their associated mental representations. Hence, differing levels of cognitive demand and/or mental states can also be represented in the form of hemodynamic response timing and intensity differences.

Resampling techniques may improve power in high-noise situations

The viability of the time-resolved fMRI analysis method presented in this study is contingent on the presence of a relatively stable degree of baseline activation within the region of interest and on the presence of events and/or tasks that induce a robust and consistent response within the brain. Care should be taken to visually inspect the averaged event-related BOLD signal time-course for each subject and experimental condition before deciding whether to proceed with the hemodynamic parameter estimation procedure. Although the shape of the underlying BOLD signal associated with visual and auditory events remained sufficiently similar to the gamma-based hemodynamic response function even at minimal stimulus intensities and stimulus durations as low as 100ms, it was apparent that the hemodynamic parameter estimates associated with such events were fairly imprecise, particularly in the temporal domain. This indicates that the perturbations in the raw, time-resolved voxel-averaged BOLD signal induced by such events might not have been clearly distinct from the inherent, stimulus-free “noise” within the visual and auditory regions of interest, leading to highly variable fit results across trials and subjects.

Conventional statistical treatments of samples involve all observations being assigned the same weight, based on the underlying assumption that every single observation is equally reliable as another. While this largely holds true for certain measures, measures of centrality in neuroimaging can be vulnerable to influence by outliers due to noise. As BOLD responses become relatively weakened at shorter event durations and lower input intensities, the signal-to-noise ratio decreases, resulting in inconsistent peak magnitude and latency estimates with large standard deviations. In Experiment 1, bootstrap resampling of the fitted hemodynamic response parameters was used in an attempt to increase the reliability of the visual and auditory hemodynamic parameter estimates associated with each level of stimulus duration and intensity. Given the assumption that the parameters of interest are normally distributed, the process of resampling the data over a large number of iterations inherently increases the weight of more consistent observations while reducing the weight of less consistent observations in the sample. This produces population-level parameter estimates, standard errors, and confidence intervals.

In Experiment 1, parameter estimates derived from a 10,000-iteration bootstrap resampling showed a reduction in the variance in peak magnitude at primary visual cortex by an approximate factor of 4 across all conditions. This reduction in variance was even more pronounced for timing-related parameters such as peak latency. There was a general reduction in the range of 95% confidence intervals of peak magnitude and peak latency estimates as a result of bootstrap resampling, improving the certainty of the estimate. However, bootstrap resampling did not meaningfully improve the parameter estimates derived for each of the trial types in Experiment 2. This can potentially be attributed to the limitations of the bootstrap resampling technique, especially in relatively small samples. Given a relatively unstable and variable dataset, the fact that bootstrap resampling reduces the weight of values further away from the estimated mean is desirable as such extreme values

are likely the result of random noise. Yet, in the case of a more reliable dataset as that obtained in Experiment 2, the inherent variability in the dataset may in fact represent meaningful differences between experimental conditions. Reducing this variability through bootstrap resampling would consequently result in an underestimation of the effects of interest. Decisions about whether to apply bootstrap resampling in time-resolved fMRI analysis should thus ideally be informed by the quality of the underlying data and parameter estimates.

Conclusion

Our study empirically demonstrated that the predicted effects of stimulus duration and intensity on evoked hemodynamic responses can be dissociated even for sub-second duration changes. For visual and auditory stimuli, duration changes affected both peak latency and peak magnitude, whereas intensity changes affected peak magnitude alone. When the underlying hemodynamic responses are robust, duration changes as small as 200 ms can result in significant peak latency changes. Careful experimental design to maximize these responses or better characterization of the HRF for brief stimulation may improve the technique. Although our study used perceptual manipulations in sensory areas, we expect that cognitive manipulations would likely have similar effects on event-related hemodynamic responses in other parts of the brain. As such, time-resolved fMRI provides a parsimonious and sound method for drawing time-based inferences about localised brain function, and it can therefore serve as an important tool for using mental chronometry in cognitive neuroscience.

References

Amunts, K., Malikovic, A., Mohlberg, H., Schormann, T., & Zilles, K. (2000). Brodmann's Areas 17 and 18 Brought into Stereotaxic Space—Where and How Variable? *NeuroImage*, *11*(1), 66–84. https://doi.org/10.1006/nimg.1999.0516

Bledowski, C. (2004). Localizing P300 Generators in Visual Target and Distractor Processing: A Combined Event-Related Potential and Functional Magnetic Resonance Imaging Study. *Journal of Neuroscience*, *24*(42), 9353–9360. https://doi.org/10.1523/JNEUROSCI.1897-04.2004

Boynton, G. M., Engel, S. A., & Heeger, D. J. (2012). Linear systems analysis of the fMRI signal. *NeuroImage*, *62*(2), 975–984. https://doi.org/10.1016/j.neuroimage.2012.01.082

Buxton, R. B., Wong, E. C., & Frank, L. R. (1998). Dynamics of blood flow and oxygenation changes during brain activation: The balloon model. *Magnetic Resonance in Medicine: Official Journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, *39*(6), 855–864.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45. https://doi.org/10.20982/tqmp.01.1.p042

Dale, A. M., & Buckner, R. L. (1997). Selective averaging of individual trials using FMRI. *NeuroImage*, *5*(4 Part 2), S47. https://doi.org/10.1016/S1053-8119(97)80006-7

Demetriou, L., Kowalczyk, O. S., Tyson, G., Bello, T., Newbould, R. D., & Wall, M. B. (2018). A comprehensive evaluation of increasing temporal resolution with multiband-accelerated protocols and effects on statistical outcome measures in fMRI. *NeuroImage*, *176*, 404–416. https://doi.org/10.1016/j.neuroimage.2018.05.011

Dux, P. E., Ivanoff, J., Asplund, C. L., & Marois, R. (2006). Isolation of a Central Bottleneck of Information Processing with Time-Resolved fMRI. *Neuron*, *52*(6), 1109–1120. https://doi.org/10.1016/j.neuron.2006.11.009

Efron, B., & Gong, G. (1983). A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation. *The American Statistician*, *37*(1), 36–48. https://doi.org/10.2307/2685844

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, *25*(4), 1325–1335. https://doi.org/10.1016/j.neuroimage.2004.12.034

Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C., Frackowiak, R. S., & Turner, R. (1995). Analysis of fMRI time-series revisited. *NeuroImage*, *2*(1), 45–53. https://doi.org/10.1006/nimg.1995.1007

Friston, K. J., Mechelli, A., Turner, R., & Price, C. J. (2000). Nonlinear Responses in fMRI: The Balloon Model, Volterra Kernels, and Other Hemodynamics. *NeuroImage*, *12*(4), 466–477. https://doi.org/10.1006/nimg.2000.0630

Goodyear, B. G., & Menon, R. S. (1998). Effect of Luminance Contrast on BOLD fMRI Response in Human Primary Visual Areas. *Journal of Neurophysiology*, *79*(4), 2204–2207. https://doi.org/10.1152/jn.1998.79.4.2204

Grinband, J., Wager, T. D., Lindquist, M., Ferrera, V. P., & Hirsch, J. (2008). Detection of time-varying signals in event-related fMRI designs. *NeuroImage*, *43*(3), 509–520. https://doi.org/10.1016/j.neuroimage.2008.07.065

Henson, R. N. A., Price, C. J., Rugg, M. D., Turner, R., & Friston, K. J. (2002). Detecting Latency Differences in Event-Related BOLD Responses: Application to Words

versus Nonwords and Initial versus Repeated Face Presentations. *NeuroImage*, *15*(1), 83–97. https://doi.org/10.1006/nimg.2001.0940

Hinton, S. C., Harrington, D. L., Binder, J. R., Durgerian, S., & Rao, S. M. (2004). Neural systems supporting timing and chronometric counting: An FMRI study. *Cognitive Brain Research*, *21*(2), 183–192. https://doi.org/10.1016/j.cogbrainres.2004.04.009

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015

Kong, R., Li, J., Orban, C., Sabuncu, M. R., Liu, H., Schaefer, A., Sun, N., Zuo, X.-N., Holmes, A. J., Eickhoff, S. B., & Yeo, B. T. T. (2019). Spatial Topography of Individual-Specific Cortical Networks Predicts Human Cognition, Personality, and Emotion. *Cerebral Cortex*, *29*(6), 2533–2551. https://doi.org/10.1093/cercor/bhy123

Lewis, L. D., Setsompop, K., Rosen, B. R., & Polimeni, J. R. (2018). Stimulus-dependent hemodynamic response timing across the human subcortical-cortical visual pathway identified through high spatiotemporal resolution 7T fMRI. *NeuroImage*, *181*, 279–291. https://doi.org/10.1016/j.neuroimage.2018.06.056

Li, J., Kong, R., Liégeois, R., Orban, C., Tan, Y., Sun, N., Holmes, A. J., Sabuncu, M. R., Ge, T., & Yeo, B. T. T. (2019). Global signal regression strengthens association between resting-state functional connectivity and behavior. *NeuroImage*, *196*, 126–141. https://doi.org/10.1016/j.neuroimage.2019.04.016

Liesefeld, H. R. (2018). Estimating the Timing of Cognitive Operations With MEG/EEG Latency Measures: A Primer, a Brief Tutorial, and an Implementation of Various Methods. *Frontiers in Neuroscience*, *12*, 765. https://doi.org/10.3389/fnins.2018.00765

Lin, F.-H., Witzel, T., Raij, T., Ahveninen, J., Wen-Kai Tsai, K., Chu, Y.-H., Chang, W.-T., Nummenmaa, A., Polimeni, J. R., Kuo, W.-J., Hsieh, J.-C., Rosen, B. R., & Belliveau, J. W. (2013). FMRI hemodynamics accurately reflects neuronal timing in the human brain measured by MEG. *NeuroImage*, *78*, 372–384. https://doi.org/10.1016/j.neuroimage.2013.04.017

Lindquist, M. A., Meng Loh, J., Atlas, L. Y., & Wager, T. D. (2009). Modeling the hemodynamic response function in fMRI: Efficiency, bias and mis-modeling. *NeuroImage*, *45*(1, Supplement 1), S187–S198. https://doi.org/10.1016/j.neuroimage.2008.10.065

Liu, H.-L., & Gao, J.-H. (2000). An investigation of the impulse functions for the nonlinear BOLD response in functional MRI. *Magnetic Resonance Imaging*, *18*(8), 931–938.

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*. https://www.frontiersin.org/articles/10.3389/fpsyg.2015.01171

Logothetis, N. K. (2003). The Underpinnings of the BOLD Functional Magnetic Resonance Imaging Signal. *Journal of Neuroscience*, *23*(10), 3963–3971.

Miezin, F. M., Maccotta, L., Ollinger, J. M., Petersen, S. E., & Buckner, R. L. (2000). Characterizing the Hemodynamic Response: Effects of Presentation Rate, Sampling Procedure, and the Possibility of Ordering Brain Activity Based on Relative Timing. *NeuroImage*, *11*(6), 735–759. https://doi.org/10.1006/nimg.2000.0568

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(24), 9868–9872.

Polimeni, J. R., & Lewis, L. D. (2021). Imaging faster neural dynamics with fast fMRI: A need for updated models of the hemodynamic response. *Progress in Neurobiology*, *207*, 102174. https://doi.org/10.1016/j.pneurobio.2021.102174

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, *84*, 320–341. https://doi.org/10.1016/j.neuroimage.2013.08.048

Stephan, K. E., Weiskopf, N., Drysdale, P. M., Robinson, P. A., & Friston, K. J. (2007). Comparing hemodynamic models with DCM. *NeuroImage*, *38*(3), 387–401. https://doi.org/10.1016/j.neuroimage.2007.07.040

Suchow, J. W., Brady, T. F., Fougnie, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. *Journal of Vision*, *13*(10), 9. https://doi.org/10.1167/13.10.9

Tombu, M., Asplund, C. L., Dux, P. E., Godwin, D., Martin, J. W., & Marois, R. (2011). A Unified attentional bottleneck in the human brain. *Proceedings of the National Academy of Sciences*, *108*(33), 13426–13431. https://doi.org/10.1073/pnas.1103583108

Vazquez, A. L., & Noll, D. C. (1998). Nonlinear aspects of the BOLD response in functional MRI. *NeuroImage*, *7*(2), 108–118. https://doi.org/10.1006/nimg.1997.0316

Wager, T. D., Vazquez, A., Hernandez, L., & Noll, D. C. (2005). Accounting for nonlinear BOLD effects in fMRI: Parameter estimates and a model for prediction in rapid event-related studies. *NeuroImage*, *25*(1), 206–218. https://doi.org/10.1016/j.neuroimage.2004.11.008

Yeşilyurt, B., Uğurbil, K., & Uludağ, K. (2008). Dynamics and nonlinearities of the BOLD response at very short stimulus durations. *Magnetic Resonance Imaging*, *26*(7), 853–862. https://doi.org/10.1016/j.mri.2008.01.008