BOLD Neurovascular Coupling Does Not Change Significantly with Normal Aging

Jack Grinband, Jason Steffener, Qolamreza R Razlighi, and Yaakov Stern

Department of Radiology, Columbia University, New York
Interdisciplinary School of Health Sciences, University of Ottawa, Ontario
Department of Neurology, Columbia University, New York
Department of Biomedical Engineering, Columbia University, New York

Abstract: Studies of cognitive function that compare the blood oxygenation level dependent (BOLD) signal across age groups often require the assumption that neurovascular coupling does not change with age. Tests of this assumption have produced mixed results regarding the strength of the coupling and its relative time course. Using deconvolution, we found that age does not have a significant effect on the time course of the hemodynamic impulse response function or on the slope of the BOLD versus stimulus duration relationship. These results suggest that in cognitive studies of healthy aging, group differences in BOLD activation are likely due to age-related changes in cognitive–neural interactions and information processing rather than to impairments in neurovascular coupling.

Key words: hemodynamic response function; HRF; aging; BOLD

INTRODUCTION

The effects of healthy aging on neurocognitive function have been widely studied using BOLD fMRI. Comparisons of neural activity between age groups often require an assumption of age equivalence in the coupling of the neural activity to the measured BOLD response [Samanez-Larkin and D’Esposito, 2008]. In the absence of this assumption, group differences could either be due to age-related changes in neurocognitive function or in neurovascular coupling. The latter explanation is underscored by the variety of morphological and vascular factors that are affected by age [Brown and Thore, 2011; Raz and Rodrigue, 2006; Raz et al., 2015]. For example, aging has been shown to be associated with vessel tortuosity [Han, 2012; Thore et al., 2007], venous callogenosis [Brown and Thore, 2011], capillary remnants [Brown, 2010], decreased vascular density [Abernethy et al., 1993], basement membrane thickening [Morris et al., 2014], decreased elasticity [Akinyemi et al., 2013], endothelial dysfunction [Hainsworth et al., 2015; Sabayan et al., 2014], and increased blood-brain barrier permeability [Heye et al., 2014; Oakley and Tharakan, 2014]. These changes in microstructure may cause reductions in perfusion [Chen et al., 2011; Fisher et al., 2013; Miners et al., 2014] and vascular reactivity [Jaruchart et al., 2015; Riecker et al., 2003; Sonntag et al., 2007], opening the possibility that age-related BOLD differences reported in fMRI studies may be due in part or wholly to changes in neurovascular coupling rather than changes in task-related neural activity.
The relationship between neural activity and the BOLD response has been shown to be approximately linear and time-invariant, and can be described by a hemodynamic impulse response function (HRF) [Boynton et al., 1996, 2012]. The HRF is a theoretical construct describing the time-dependent changes in the BOLD signal to an infinitesimally small neural input, and is often approximated by the empirically measured hemodynamic response (HDR) to a neural input of short, but finite, duration. Though the vascular mechanisms that produce the specific shape of the HRF are not fully understood, it is reasonable to hypothesize that age-related changes in the cerebral vasculature may have a direct effect on both the shape and amplitude of the HRF. However, studies of the effect of age on the shape of the HRF using estimates of the HDR have produced mixed results. For example, Richter and Richter [2003] and Aizenstein et al. [2004] found age-related reductions in the magnitude of the undershoot. Taoka et al. [1998] and Handwerker et al. [2007] found an age-related delay in the time to peak, whereas Huetell et al. [2001] found an age-related acceleration in time to peak. In contrast, D’Esposito et al. [1999] found no significant age differences in the HRF shape, and Ward et al. [2008] argued that the HRF does not change with age.

Effects of age on the strength of BOLD neurovascular coupling have also been mixed. Age-related decreases in the BOLD response have been found in the motor [Buckner et al., 2000; Hesselmann et al., 2001; Riecker et al., 2003; Tekes et al., 2005] and the visual [Ances et al., 2009; Fabiani et al., 2014; Handwerker et al., 2007; Ross et al., 1997; Tekes et al., 2005; Ward et al., 2015] cortices, suggesting reductions in coupling strength. In contrast, a number of studies have found no age-dependent differences (Aizenstein et al., 2004; Brodtmann et al., 2003; D’Esposito et al., 1999; Huettel et al., 2001; Richter and Richter, 2003; Rosengarten et al., 2003; Schröter et al., 2004). In addition, Mehagnoul-Schipper [2002] showed no age differences in the BOLD response using fMRI, but decreases using fNIRS, and Mattay et al. [2002] found a BOLD increase, which was argued to be compensatory. In sum, there does not appear to be a consensus on the effects of healthy aging on either the shape of the time course or the strength of coupling.

**HRF Estimation**

The estimation of the HRF in aging studies is typically done using either deconvolution or event-triggered averaging [Glover, 1999; Josephs et al., 1997]. However, estimation of the HRF using deconvolution can be problematic, especially if the stimulus or response duration is constant across trials. Specifically, when the HRF estimate relies exclusively on repeated presentations of very brief (<500 ms) sensory stimuli or single button presses, the nonlinear component of the HRF [Birn et al., 2001; Glover, 1999; Yesilyurt et al., 2008] can become a significant source of variance and may bias the deconvolved shape to be more representative of longer neural inputs. Nonlinearities also exist at very long stimulus durations [Janz et al., 2000; Martindale et al., 2005] and long stimulation blocks are associated with both neural and vascular attenuation [Obrig et al., 2002], which can affect the interpretation of neurovascular coupling in block designs. Furthermore, constant duration events have lower entropy than jittered events and, thus, may not be optimal for estimating coupling strength [Liu, 2004; Liu and Frank, 2004; Liu et al., 2001; Miezin et al., 2000]. This can be particularly problematic since aging is associated with greater levels of noise in the BOLD data [D’Esposito et al., 1999; Huettel et al., 2001] thus increasing the likelihood of mis-specifying the shape of the impulse response.

Estimation of the HRF using event-triggered averaging requires the assumption that the neural activity associated with the sensory stimulus or with the motor response is infinitely small (an impulse) and does not contribute significantly to the shape or amplitude of the HRF. However, this assumption may not hold, particularly if elderly subjects use different behavioral strategies than younger subjects, such as attending more/less intently to a visual stimulus or pressing a button with more/less force. If disparate behavioral strategies are used, then differences in underlying neural processing will bias the HRF shape estimate, as well as any estimate of BOLD magnitude, i.e., $HDR = HRF \otimes n(t)$, where $n(t)$ represents the duration of neural processing that may vary as a function of behavioral strategy.

Experimental paradigms for studying neurovascular coupling typically rely on very simple behaviors, such as finger tapping or passive viewing of flashing checkerboards that are intended to minimize behavioral differences between age groups. However, even for such simple tasks, behavior can differ significantly as a function of age. For example, the magnitude of the BOLD response in the primary motor cortex is sensitive to the force, duration, and speed of the motor response [Peck et al., 2001; Rao et al., 1996; Sadato et al., 1997; Ward et al., 2008] which have been shown to decline with age [Schmidt and Lee, 2011]. Thus, behavioral parameters of the movement, rather than neurovascular effects, could contribute to apparent age-related influences on the shape and amplitude of the estimated motor HRFs.

Behavioral factors can also affect the amplitude of sensory responses. For example, attention is known to modulate primary sensory activity [Boynton, 2011], but because vigilance/sustained attention degrades with age [Berardi et al., 2001; Brache et al., 2010; McVay et al., 2013; Mouloua and Parasuraman, 1995; Parasuraman et al., 1989; Staub et al., 2013, 2014a,b,c; Thompson, 2014], BOLD differences could be due to unequal attentional loads between groups. Such attentional differences can be exacerbated in experimental paradigms that use long (15–30 s) stimulus durations or long intertrial intervals (ITI) to estimate neurovascular coupling [Brodtmann et al., 2003; Buckner et al., 2000; D’Esposito et al., 2004].
et al., 1999; Fabiani et al., 2014; Huettel et al., 2001; Richter and Richter, 2003; Schroeter et al., 2004; Tekes et al., 2005; Ward et al., 2015]. Furthermore, even when attentional demands are nominally equal, age-related decrements in the visual system, such as increases in visual detection and discrimination thresholds [Owsley, 2011], can lead to compensatory behavioral adaptations such as longer behavioral response times, with corresponding changes in BOLD activity [Carp et al., 2012; Yarkoni et al., 2009].

The goal of this study was to minimize these potentially confounding variables and determine how the time course and strength of neurovascular coupling changes with age. We chose to use time-varying visual and auditory stimulation in the context of a rapid event-related fMRI design [Grinband et al., 2008]. Using multiple durations of the sensory stimuli in the range (500–4000 ms) allowed us to minimize the effect of BOLD nonlinearities that dominate at durations less than 500 ms and produce an experimental design with greater signal entropy compared to constant duration designs, while providing an efficient means of estimating the neurovascular coupling strength. Furthermore, in contrast to the use of motor responses, it is easier to control sensory stimulus durations, obviating the need to monitor motor-related behavioral parameters, such as force magnitude, speed, and duration [Peck et al., 2001; Rao et al., 1996; Sadato et al., 1997; Ward et al., 2008], which are known to vary with age [Schmidt and Lee, 2011]. Finally, rapid event-related designs are less sensitive to attention and expectation effects than highly regular, block or slow event-related designs [Josephs et al., 1997]. By systematically varying the duration of the sensory input, it is possible to use deconvolution to simultaneously characterize the shape of the HRF and estimate the neurovascular coupling strength, while minimizing the effects of nonlinearities present at brief stimulus durations. We were able to dissociate neurovascular from neural effects and demonstrate that healthy aging does not have a physiologically meaningful effect on the time course and strength of neurovascular coupling.

METHODS

Subjects

Informed consent, as approved by the Internal Review Board of the College of Physicians and Surgeons of Columbia University, was obtained prior to study participation, and after the nature and risks of the study were explained. Participants were paid for their participation in the study. Healthy subjects, 55 young (22M/33F) and 34 elderly (14M/20F) were recruited from the local community. The age range for young subjects was 18–30 (mean = 24.8, SD = 2.7); the age range for older subjects was 54–74 (mean = 65.1, SD = 4.6). In the elderly group, 10 individuals had managed hypertension, 3 had managed endocrine disease, and 2 were active smokers; 5 elderly subjects did not report on these medical conditions.

Participants were recruited using established market mailing procedures to equalize the recruitment procedures of young and elderly. Participants who responded to the mailing were telephone screened to ensure that they met basic inclusion criteria (right handed, English speaking, no psychiatric or neurological disorders, normal, or corrected-to-normal vision). Individuals that passed the telephone screen were further screened in person and a Mattis Dementia Rating Scale score of at least 136 was required for retention in the study. Exclusion criteria included visual deficits such as retinal degeneration, cataracts, glaucoma, corneal disorders, night blindness, and corrected visual acuity lower than 20/20; uncontrolled high blood pressure (systolic >180 mm Hg or diastolic >105 mm Hg); active hepatic disease or primary renal disease requiring dialysis, primary untreated endocrine disease (e.g., insulin dependent diabetes Type I or II), medications that target CNS (e.g., neuroleptics, anticonvulsants, antidepressants, or benzodiazepines); history of psychosis or electroconvulsive therapy; major depressive disorder, bipolar disorder, anxiety disorder, diagnosed learning disability, dyslexia, or ADHD; history of alcohol or drug abuse dependence; any brain disorders such as stroke, infarct, subcortical lacunae, tumor, infection, epilepsy, multiple sclerosis, degenerative diseases, head injury (LOC > 5 min), or any other abnormalities as judged by a radiologist (periventricular caps or small white matter hyperintensities were not excluded).

Behavioral Task

Subjects participated in one or two audiovisual tasks (Task 1: attend visual/ignore auditory; Task 2: attend auditory/ignore visual). Each task was 6-min long and part of a battery of tasks administered during the experimental session. The visual stimulus consisted of a reversing checkerboard presented at 6 Hz. The auditory stimulus consisted of two tones (600 and 1000 Hz) alternating at 10 Hz. The rapid alternation of the auditory stimuli created a percept of a quavering/vibratory nature. Both the visual and auditory stimuli were perceived as distinct and highly salient integrated objects. Furthermore, the duration of each stimulus was rounded such that it was a multiple of the duration of an individual checkerboard or tone, in order to eliminate visual screen “tearing” or auditory artifacts at stimulus termination, respectively. Subjects were instructed to either (1) attend to the visual stimulus and press a button as quickly as possible after the checkerboard disappeared, while ignoring the auditory stimulus or (2) attend to the auditory stimulus and press a button as quickly as possible after the tones disappeared, while ignoring the visual stimulus. The button press was used to maintain focused attention on the target stimulus. The instructions to attend to only one of the stimuli allowed us to compare the shape and amplitude of the BOLD response under two different attentional loads. Stimulus durations were randomly chosen from a uniform...
distribution to minimize expectation effects (visual stimuli = 0.5–4.0 s; auditory stimuli = 0.60–4.0 s). The ITIs were also randomly chosen from a uniform distribution (ITI = 2.5–9.0 s). Since previous work has shown that the BOLD response has a refractory period [Huetel and McCarthy, 2000, 2001], the randomization of ITI durations distributed the effects of any nonlinearities equally across the sensory stimuli. For each subject the correlation between the onsets of the visual and auditory stimuli was near zero. The visual stimulus was back projected onto a screen using an LCD projector. The auditory stimulus was presented using MR compatible headphones. The task was administered on a Macbook Pro using the Psychophysics Toolbox [Brainard, 1997; Pelli, 1997] and Matlab (www.mathworks.com).

Assumptions

To estimate the shape of the hemodynamic impulse response function, the neural input function must be known. The BOLD response most closely reflects local field potentials [Lauritzen and Gold, 2003; Logothetis, 2008; Logothetis et al., 2001] and in animals, this input function can be measured invasively using electrophysiology [Logothetis, 2008; Murayama et al., 2010] and optical imaging [Hillman, 2014; Iordanova et al., 2015]. In humans, methods exist for simultaneous measurement of field potentials and hemodynamic activity, though they require neurosurgical interventions [Keller et al., 2009; Niwayama and Yamakawa, 2014]. Thus the input function is usually not known. However, physiological work has demonstrated that in the primary sensory cortex, the time course of the neuronal response closely matches the time course of the sensory stimulus [Albrecht and Hamilton, 1982; Boynton, 2011; Carandini and Sengpiel, 2004; Logothetis et al., 2001; Zhan et al., 2005; Zhang et al., 2007]. Thus, we assume that within primary visual and primary auditory cortex,

\[ n(t) \approx \text{sensory input} \]

where \( n(t) \) represents the time course of the neural input function. The hemodynamic response can then be approximated by

\[ \text{HDR} = \text{HRF} \otimes n(t). \quad (1) \]

The hemodynamic response is also related to the neurovascular coupling strength \( (\alpha) \), that is, the scaling factor for a steady state input, and the intensity of the neural input \( (\beta) \), such that

\[ \text{HDR} = a\text{HRF} \otimes \beta n(t), \quad (2) \]

where HRF and \( n(t) \) are normalized to 1. Normalization consists of dividing the HRF by the integral of the HRF and dividing \( n(t) \) by the integral of \( n(t) \), effectively dissociating the shape information (HRF and \( n(t) \)) from the amplitude of the response \( (\alpha \) and \( \beta \)). This is a linear equation, \( y = bx \), in which, \( y = \text{HDR} \), \( x = \text{HRF} \otimes n(t) \), and \( b = \alpha \beta \). Moreover, it demonstrates that the slope of the function, BOLD magnitude (HDR) versus the neural input \( (n(t)) \), is proportional to both \( \alpha \) and \( \beta \). To disambiguate \( \alpha \) and \( \beta \), we assume that \( \alpha \) is not sensitive to cognitive effects and is constant over the range of brief stimulus durations (500–4000 ms) used in our study. In contrast, the intensity of the neural response, \( \beta \), has been reliably shown to be sensitive to cognitive effects—specifically relevant to this study is the attentional modulation of neural activity [Boynton, 2011; Grady et al., 1997; Kastner et al., 1999; Luck et al., 1997; Watanabe et al., 2011; Woldorff et al., 1993].

Thus, by manipulating the duration of the sensory stimulus and attentional load, it is possible to dissociate age-related effects on \( \alpha \) and \( \beta \). Specifically, if a difference in \( \beta \) due to task-related attentional modulation is detectable, then a comparable age-related difference in \( \alpha \) is also detectable, if such difference existed.

MRI Data Acquisition

Imaging was performed on a 3T Philips Achieva Scanner using an 8-channel SENSE head coil. Structural scans were performed using the 3D MPRAGE sequence (165 slices; 256 × 256; FOV = 256 mm; echo time = 3 ms; TR = 25 ms; flip angle = 45°). Functional scans used EPI-BOLD (TE = 20 ms; TR = 2000 ms; 41 slices; 112 × 112; FOV = 224 mm; voxel size = 2 mm × 2 mm × 3 mm) for 15 young and 14 old subjects and (TE = 20 ms; TR = 1,000 ms; 22 slices; 80 × 80; FOV = 240 mm; voxel size = 3 mm × 3 mm × 5.5 mm) for the remaining subjects.

Data Analysis

All image analysis was done using the FMRIB Software Library [FSL; Jenkinson et al., 2012] and Matlab (Mathworks, Natick, MA; www.mathworks.com). The data were motion-corrected (FSL-MCFLIRT), slice-time corrected, high-pass filtered (at 0.02 Hz, 50 s), and spatially smoothed (full width at half maximum = 5 mm) and Melodic independent component analysis was used to remove motion and scanner-related artifacts in the data [Beckmann and Smith, 2004]. Standard statistical parametric mapping was performed in original T2* space.

To identify the voxels in the visual or auditory cortices of each individual subject to use in subsequent analyses, we constructed a design matrix in which each regressor matched the time course (onset and duration) of the stimulus, and each stimulus modality was modeled as a separate regressor. No assumptions about the shape of the HRF were made other than that it be physiologically plausible—FSL-FLOBS was used to provide a flexible convolution basis set [Woolrich et al., 2004]. In brief, the FLOBS basis set was created by first parameterizing the HRF into four half-period cosines with six parameters. Then singular value decomposition was performed on 1,000 probabilistically
generated HRFs to create a set of eigen-HRFs/basis functions. The first three eigen-HRFs are very similar to commonly used canonical, delay (temporal) derivative, and width (dispersion) derivative of a Gamma/Gaussian parameterized HRF. We used these three orthogonal basis functions to convolve each subject's sensory regressors (i.e., the basis set was identical for all subjects in the study).

The design matrix contained a total of 6 regressors (3 visual, 3 auditory). An F-test on the visual regressors and an F-test on the auditory regressors was performed using a voxel-wise threshold of \( P = 0.05 \) and a cluster threshold of \( P = 0.05 \) using Gaussian random field theory. The thresholded map was then multiplied either by a mask of the occipital cortex or of the superior temporal cortex, respectively, and binarized. Both masks were originally defined in standard MNI152 space and transformed to each individual's T2* space. The resulting voxels were labeled as “visual” or “auditory” and used in all subsequent analyses.

The shape of the HRF was estimated by averaging the parameter estimates across all visual or all auditory voxels and performing a weighted average using the FLOBS basis functions, i.e. \( HRF = pe_1 \cdot f_1 + pe_2 \cdot f_2 + pe_3 \cdot f_3 \), where \( pe_x \) is the mean parameter estimate across voxels and \( f_x \) is the basis function. HRFs with parameter estimates that exceeded three standard deviations from the mean were excluded from the analysis.

**Figure 1.**
Impulse response functions. (A) HRFs were plotted for young subjects from the visual and auditory cortex. Gray lines represent data from individual subjects; thick solid lines represent the group mean. (B) HRFs from old subjects. (C) Mean responses from the two groups with standard errors are plotted together. [Color figure can be viewed at wileyonlinelibrary.com]

**Testing for Equivalence in Shape**

We assessed shape equivalence in two ways: (1) quantifying shape features and (2) performing a variance analysis to determine whether any significant differences, if detected, are physiologically meaningful. To test for equivalence, we compared groups on time to peak, time to trough, trough magnitude, rise slope, and fall slope, and performed a power analysis to quantify the number of subjects needed to detect significant differences in each feature. We assumed that most fMRI studies of cognitive aging use approximately 15–30 subjects per group. If the number of subjects needed to detect HRF shape differences dramatically exceeds this number, then any existing HRF shape differences can be considered not to be physiologically meaningful. We performed a power analysis (Fig. 1D) in which the \( t \) statistic is defined as:

\[
t = \frac{|\Delta HRF|}{\sqrt{\sigma^2_{young} + \sigma^2_{old}}}
\]
where \( \Delta \text{HRF} \) is the difference between an HRF feature between young and old and \( \sigma^2 \) is the group variance. The test was assumed to be two-tailed with an equal number of subjects across groups, with alpha = 0.05, and beta = 0.80.

A second test of equivalence was performed using a variance analysis in order to compare between-group differences in HRF shape to within-group individual differences in HRF shape. We assumed (1) that to be meaningful, any group differences in HRF shape must be equal to or larger than the individual natural variation in HRF shape between subjects and (2) that any variation in HRF shape that is smaller than within-group differences will not have significant influence on the vast majority of cognitive aging studies. We computed the sum of squared errors between the mean HRFs for the two groups and compared this value to (1) the sum of squared errors between the mean young HRF and individual young HRFs and (2) the sum of squared errors between the mean old HRF and individual old HRFs. A \( t \)-test was used to test whether the sum of squared errors between groups was larger than the error within group. This analysis was repeated using the canonical HRF (cHRF) [Friston et al., 1994; Josephs et al., 1997] to determine whether the sum of squared errors between the cHRF and our sample mean HRF was bigger than the sum of squared errors between the two age groups. The cHRF was represented by the difference of two gamma functions \((\Gamma(6,1) - \Gamma(16,1)/6)\).

### Testing for Equivalence in Amplitude

A whole brain regression analysis was performed to identify regions in visual and auditory cortex that responds to attentional effects. The model consisted of one visual and one auditory regressor, composed of boxcars with onset and offset times matching the sensory stimuli. Mixed-effects group analysis was performed at the second level (voxel-wise threshold, \( P = 0.01 \); cluster threshold, \( P = 0.05 \)). A contrast of attended greater than unattended for each stimulus condition was used to identify regions modified by attention. These regions were then intersected with primary sensory cortex and binarized to create masks for testing whether the strength of neurovascular coupling changes with age.

To quantify amplitude, a regression analysis was performed in which the sensory stimuli were grouped into five temporal quantiles (i.e., each group of stimuli consisted of short, medium-short, medium, medium-long, or long durations). For each quantile of trials, a regressor was created in which each trial was represented by a boxcar with duration equal to the mean of the quantile and normalized to the integral (i.e., integral of each boxcar = 1). Then each regressor was convolved with the subject-specific, regional HRF. The resulting design matrix consisted of 10 regressors (5 visual and 5 auditory). The parameter estimates for each regressor were averaged across subjects and plotted as a function of mean stimulus duration. The slope of this function was assumed to represent \( \beta \) and was computed using linear regression.

To determine whether any differences in BOLD magnitude, if they existed, were physiologically meaningful, we compared any differences in slope between age groups due to changing in neurovascular coupling (\( \alpha \)) to differences in slope due to neural intensity (\( \beta \)). Using the attentional comparison served two purposes. First, since it is well established that attention can modulate neuronal activity, manipulating the subject’s attentional state demonstrates whether a small difference in neural firing can result in a detectable change in slope. Second, we assumed that, to be physiologically meaningful, differences in neurovascular coupling between age groups must be equal to or larger than slope differences generated by task-related behavioral variables (e.g., attention).

We also tested whether the magnitude of the BOLD response is linear within the range of temporal values in our experimental design. This was done using a step-wise regression, such that for each subject, the linear component was removed and a quadratic or exponential model was tested on the residual.

### RESULTS

#### Shape Differences

To determine whether aging affects the shape of the HRF, we estimated the HRF in the visual and auditory cortices for the group of young (Fig. 2A) and elderly (Fig. 2B) subjects. A comparison of group means (Fig. 2C) indicates that young and old HRFs are very similar. To quantify the similarity, we computed commonly measured features of the HRF shape, i.e., time to peak, time to trough, trough magnitude, rising slope, and falling slope. No significant differences (Student \( t \)-test, \( P < 0.05 \)) were detected between groups for either visual (Fig. 1A) or auditory (Fig. 1B) HRF features. A power analysis was performed to determine the minimum number of subjects needed to find a significant difference for each feature (Table I).

As the number of subjects in the sample approaches the number of subjects in the population, population differences will ultimately be detected. To determine whether such differences could be physiologically meaningful when making group comparisons, we compared between-group variability versus the within-group variability in HRF shape. Individual differences in the sum of squared errors between individual HRFs and the group mean are more than \( \sim 3 \) times larger than the sum of squared errors between the young and old group means (Fig. 3; young vis SSE > young-old vis SSE, \( P < 4 \times 10^{-18} \); old vis SSE > young-old vis SSE, \( P < 9 \times 10^{-8} \); young aud SSE > young-old aud SSE, \( P < 3 \times 10^{-13} \); old vis SSE > young-old SSE, \( P < 1 \times 10^{-4} \)).
We further compared young and old group means with the cHRF that is used in most neuroimaging studies of cognitive aging. Figure 4A shows that the cHRF has a slower time to peak and trough than either the young or old group mean. The double gamma fit to the mean (Supporting Information Fig. S1) across all HRFs was \((C(6.66,1) \cdot C(6.77,1)/1.05\). Furthermore, the misalignment between the cHRF and the group means (Fig. 4B, black bar) results in error comparable to those from individual differences (red and blue bars). In fact, the error generated by using the cHRF is much larger than the error between young and old group means (3.6 times larger for visual and 4.6 times larger for auditory cortex).

**Magnitude Differences**

Magnitude differences between groups and conditions can be detected by measuring the slope of the BOLD magnitude versus stimulus duration function. To test whether changes in slope are detectable, we used attention to manipulate \(\beta\), the slope due to the intensity of the neuronal input function. Voxels sensitive to attentional effects

---

**TABLE I. Group differences in HRF shape and power analysis**

<table>
<thead>
<tr>
<th>HRF feature</th>
<th>Young</th>
<th>Old</th>
<th>(P) value</th>
<th>Number of subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to peak visual (s)</td>
<td>4.7 (0.10)</td>
<td>4.9 (0.12)</td>
<td>0.289</td>
<td>282</td>
</tr>
<tr>
<td>Time to peak auditory (s)</td>
<td>4.6 (0.09)</td>
<td>4.8 (0.16)</td>
<td>0.308</td>
<td>239</td>
</tr>
<tr>
<td>Time to trough visual (s)</td>
<td>10.6 (0.19)</td>
<td>11.3 (0.36)</td>
<td>0.101</td>
<td>140</td>
</tr>
<tr>
<td>Time to trough auditory (s)</td>
<td>9.9 (0.18)</td>
<td>10.6 (0.40)</td>
<td>0.101</td>
<td>105</td>
</tr>
<tr>
<td>Trough mag visual</td>
<td>−0.138 (0.02)</td>
<td>−0.096 (0.03)</td>
<td>0.16</td>
<td>183</td>
</tr>
<tr>
<td>Trough mag auditory</td>
<td>−0.179 (0.02)</td>
<td>−0.138 (0.05)</td>
<td>0.427</td>
<td>406</td>
</tr>
<tr>
<td>Rise slope visual</td>
<td>0.02 (2 \times 10^{-4})</td>
<td>0.02 (2 \times 10^{-4})</td>
<td>0.838</td>
<td>14938</td>
</tr>
<tr>
<td>Rise slope auditory</td>
<td>0.02 (3 \times 10^{-4})</td>
<td>0.02 (3 \times 10^{-4})</td>
<td>0.338</td>
<td>588</td>
</tr>
<tr>
<td>Fall slope visual</td>
<td>−0.02 (3 \times 10^{-4})</td>
<td>−0.02 (4 \times 10^{-4})</td>
<td>0.449</td>
<td>1091</td>
</tr>
<tr>
<td>Fall slope auditory</td>
<td>−0.02 (4 \times 10^{-4})</td>
<td>−0.02 (5 \times 10^{-4})</td>
<td>0.649</td>
<td>2609</td>
</tr>
</tbody>
</table>

Group means for each HRF feature are listed with standard errors. A two-sided \(t\)-test was performed. None of the \(P\) values exceeded \(P < 0.05\) (no correction for multiple comparisons). A power analysis with alpha = 0.05, beta = 0.80, and pooled variance indicates the number of subjects needed per group to detect a significant difference between groups. The number of subjects needed per group exceeds typical sample sizes in fMRI cognitive aging studies.
were identified by performing a whole brain analysis to detect regions in visual and auditory cortex that were activated by the task (Fig. 5A) and that showed greater activation in the attended than unattended condition (Fig. 5B). No attentional effect was detected in primary visual cortex, however, voxels in auditory cortex showed bilateral attentional modulation. These voxels were then used to create a mask for comparing attentional and age-related effects on slope. The slope was significantly greater in the attended than unattended condition for both young and old subjects (Fig. 6A). However, no significant slope differences were detected between young and old subjects (Fig. 6B). A power analysis demonstrated that the number of subjects necessary to detect a group difference in magnitude dramatically exceeds the number needed to detect task-related effects (Table II).

The comparison of slopes assumes that the BOLD response is linear across the stimulus duration used in the

Figure 3.
Sum of squared errors. The SSE between individual young HRFs and the young group mean (red) and between individual old HRFs and the old group mean (blue) is more than three times larger than the SSE between the group means. Any differences in shape due to age are dwarfed by individual differences between subjects for both visual and auditory HRFs. Error bars represent standard error. [Color figure can be viewed at wileyonlinelibrary.com]

Figure 4.
Comparison to cHRF. (A) Both young and old mean HRFs are faster than the cHRF. (B) The sum of squared errors was computed between the cHRF and each young subject (red), between cHRF and each old subject (blue), between the cHRF and the sample mean (i.e., the mean of young and old HRFs; black), and between the young and old means (purple). The error generated by using the cHRF is 3.6 times larger than the difference between young and old HRFs in visual cortex and 4.6 times larger in auditory cortex. [Color figure can be viewed at wileyonlinelibrary.com]
study. To explicitly test this assumption, we performed a step-wise regression to remove the linear variance from the data. We then tested whether the residual contained a significant quadratic or exponential relationship to stimulus duration. The quadratic (Supporting Information Fig. S1) and exponential (Supporting Information Fig. S2) terms were not significantly different from zero, nor did they differ between attentional load or age groups, suggesting that the BOLD relationship for the stimulus parameters used in the study was linear.

DISCUSSION

Though it is possible to model individual differences in the shape of the hemodynamic impulse response function, most neuroimaging studies use the chRF as the convolution kernel for modeling BOLD data [Grinband et al., 2008; Handwerker et al., 2004]. This assumption implies that the hemodynamic impulse response does not vary across subjects or brain regions, i.e., that the HRF is age-invariant. Nevertheless, it is common in cognitive aging studies for the conclusions about task-evoked brain activity to be qualified by the possibility that any age-related effects may in fact be due to differences in neurovascular coupling rather than the underlying neural response, i.e., that the HRF is not age-invariant. The validity of the assumption about the effects of age on the HRF shape and the strength of neurovascular coupling has not been conclusively resolved, creating uncertainty in the cognitive aging neuroimaging community.

Previous work has demonstrated significant variability in the shape of the HRF between individuals [Handwerker et al., 2012; Handwerker et al., 2004], and some individual differences in shape could reflect age-related changes in vascular function [Brown and Thore, 2011; Raz and Rodrigue, 2006; Raz et al., 2015]. In fact, a number of studies have demonstrated abnormalities in the impulse response, such as reductions in the size of the undershoot [Aizenstein et al., 2004; Richter and Richter, 2003] and shifts in the time to peak [Handwerker et al., 2007; Huettel et al., 2001; Taoka et al., 1998]. Two common features of these studies are that the temporal structure of the task did not vary between trials, i.e., each BOLD response either reflected a regular, constant duration event (200–1500 ms) [Aizenstein et al., 2004; Buckner et al., 2000;
D’Esposito et al., 1999; Handwerker et al., 2007; Huettel et al., 2001] or a regular, constant duration block of events (20–30 s) [Ances et al., 2009; Brodtmann et al., 2003; Fabiani et al., 2014; Hesselmann et al., 2001; Mehagnoul-Schipper et al., 2002; Richter and Richter, 2003; Ross et al., 1997; Schroeter et al., 2004; Taoka et al., 1998; Tekes et al., 2005; Ward et al., 2015].

A lack of temporal variance in the task design can result in several disadvantages for estimating HRF shape. First, nonlinearities at very brief [Huettel, 2004] or very long durations [Janz et al., 2000; Martindale et al., 2005; Obrig et al., 2002] of neuronal activity can bias the estimate of the HRF. Second, if age-related cognitive changes affect the duration of neural processing, it could shift the BOLD response out of the linear range, but for only one group. Third, behavioral differences (e.g., force, speed, and/or duration) that are sensitive to age, unless explicitly modeled using different neural input functions, can influence the estimate of the motor HRF. A similar effect can occur due to age-related differences in vigilance, which can influence assessments of neurovascular coupling for tasks with long stimulus durations or ITIs.

**Figure 6.**

Slope differences due to attention and age. (A) Attention modulates neural activity in auditory cortex causing a change in the slope of the BOLD signal versus stimulus duration. Attended stimuli were associated with larger BOLD responses indicating that changes in $\beta$ are detectable. (B) No significant difference in slope was present between groups suggesting that neurovascular coupling ($\alpha$) does not change with age. [Color figure can be viewed at wileyonlinelibrary.com]

**TABLE II. Group difference in response magnitude and power analysis**

<table>
<thead>
<tr>
<th></th>
<th>Attended slope</th>
<th>Unattended slope</th>
<th>P value</th>
<th># of subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young</td>
<td>53.8 (5.3)</td>
<td>42.9 (3.4)</td>
<td>0.047</td>
<td>86</td>
</tr>
<tr>
<td>Old</td>
<td>48.7 (8.5)</td>
<td>25.7 (4.6)</td>
<td>0.015</td>
<td>19</td>
</tr>
</tbody>
</table>

A t-test (one-sided for attention, two-sided for age) was performed to detect group differences in the slope of the parameter estimate versus stimulus duration function. A power analysis with $\alpha = 0.05$, $\beta = 0.80$, and pooled variance indicates the number of subjects needed per group to detect a significant difference in neurovascular coupling strength between groups greatly exceeds the number needed to detect attentional effects and exceeds typical sample sizes in fMRI cognitive aging studies.

D’Esposito et al., 1999; Handwerker et al., 2007; Huettel et al., 2001] or a regular, constant duration block of events (20–30 s) [Ances et al., 2009; Brodtmann et al., 2003; Fabiani et al., 2014; Hesselmann et al., 2001; Mehagnoul-Schipper et al., 2002; Richter and Richter, 2003; Ross et al., 1997; Schroeter et al., 2004; Taoka et al., 1998; Tekes et al., 2005; Ward et al., 2015].

A lack of temporal variance in the task design can result in several disadvantages for estimating HRF shape. First, nonlinearities at very brief [Huettel, 2004] or very long durations [Janz et al., 2000; Martindale et al., 2005; Obrig et al., 2002] of neuronal activity can bias the estimate of the HRF. Second, if age-related cognitive changes affect the duration of neural processing, it could shift the BOLD response out of the linear range, but for only one group. Third, behavioral differences (e.g., force, speed, and/or duration) that are sensitive to age, unless explicitly modeled using different neural input functions, can influence the estimate of the motor HRF. A similar effect can occur due to age-related differences in vigilance, which can influence assessments of neurovascular coupling for tasks with long stimulus durations or ITIs.
In this study, the influence of physiological noise and nonlinear hemodynamic effects were minimized by using sensory stimuli that elicit a reliable response in sensory cortex and using stimulus durations that sampled the linear range of the BOLD response. Furthermore, by using a rapid event-related design, we minimized attentional differences between groups that may affect estimates of neurovascular coupling. Our data showed no significant difference between the shapes of young and old HRFs, indicating that the same HRF can be used when comparing groups in cognitive aging fMRI studies. Thus, the decision to use the cHRF for both young and old subjects, as is commonly done, will not result in systematic biases in group comparisons. However, it will result in lower signal to noise than using individualized HRFs [Handwerker et al., 2004] with model errors substantially larger than any due to group differences in HRF shape (Fig. 5B). Furthermore, our results suggest that if individualized HRFs are not available, a slightly faster version of the canonical HRF should be used with parameters of (\(\Gamma(6.66,1)\) – \(\Gamma(6.77,1)/1.05\), rather than the traditional \(\Gamma(6,1)\) – \(\Gamma(16,1)/6\). Finally, we found no difference in the slope of the BOLD versus stimulus duration function, suggesting that the strength of neurovascular coupling (\(\alpha\)) does not change with normal aging. The lack of a difference cannot be explained by insensitivity of the measurement since a significant reduction in the slope (\(\beta\)) was clearly detectable with attentional modulation.

There are several caveats to this study. First, the neural input function was not measured directly. Thus, our conclusions only apply to primary sensory cortices where the neural input function can be assumed to match the time course of the stimulus [Albrecht and Hamilton, 1982; Boynton, 2011; Carandini and Sengpiel, 2004; Logothetis et al., 2001; Zhan et al., 2005; Zhang et al., 2007]. Though aging has been shown to affect the firing properties of sensory neurons [Fu et al., 2013; Leventhal et al., 2003; Schmolesky et al., 2000; Wang et al., 2005; Zhang et al., 2008], its effect on aggregate neuronal activity in sensory cortex is relatively small. For example, Celesia and Daly [1977] showed that the P1 and N1 visual evoked potentials are delayed by only \(\sim 1.7\) ms/decade. Cepoiene et al. [2008] found a 13 ms aging-related delay in P1 and no significant delay in N1, and Fabiani et al. [2014] showed a 20 ms delay in C1. In the auditory system, simple tones do not elicit any age-related delays in P1 and N1 onset or peaks [Cepoiene et al., 2008; Tremblay and Ross, 2007; Tremblay et al., 2004]. Thus, it is unlikely that age-related conduction delays could influence the shape of the HRF in visual or auditory cortex. The neural input function in higher order cognitive regions is presumably different from sensory cortex, thus, estimating the shape of the HRF outside nonprimary brain regions may not be feasible using our method.

Second, our elderly subjects were extensively assessed for a variety of general and brain-specific pathologies. A review of previous studies [Aizenstein et al., 2004; Ances et al., 2009; Buckner et al., 2000; D’Esposito et al., 1999; Handwerker et al., 2007; Hesselmann et al., 2001; Huettel et al., 2001; Mehagnoul-Schipper et al., 2002; Richter and Richter, 2003; Ross et al., 1997; Taoka et al., 1998; Tekes et al., 2005], showed a variety of exclusion criteria, ranging from (1) neurologic, radiologic, and general medical criteria to (2) only neurologic criteria to (3) no clinical criteria. This raises the question of whether or not some previous studies found age-related changes in neurovascular coupling that were caused by clinical factors rather than healthy aging.

Third, our sample of elderly subjects was limited to individuals under 75 years old. Thus, over a range of \(\sim 40\) years, there are no significant differences in the HRF. It is possible that at the tails of the distribution, for example, 75–100 years old, the HRF may, in fact, be sensitive to age. However, because of comorbidities, it is difficult to be confident that any detectable difference is a result of healthy aging and not due to a disease process.

Fourth, our results apply primarily to BOLD data. Other physiological measures that quantify cerebral blood flow, cerebral blood volume, calcium, nitric oxide, etc. can also be used to characterize neurovascular coupling. Moreover, age-equivalence in the BOLD HRF does not imply age-equivalence in other processes. For example, Ances et al. [2009] showed that age does not have the same effects on CBF and BOLD during task-evoked responses and hypercapnia challenges. Similarly, Fabiani et al. [2014] demonstrated differential effects of aging on oxy- and deoxy-hemoglobin.

Finally, it is impossible to state with certainty that group differences do not exist in the HRF. In the limit (i.e., if we were able to sample the entire population), there would undoubtedly be a detectable group difference. However, we showed that any group differences in the HRF that do exist are unlikely to have meaningful physiological effects that exceed the within-group variability. In fact, using the cHRF, and thus discarding individual differences in HRF shape, results in model error that is \(\sim 4\) times larger than the difference in shape between young and old subjects.

In sum, our results indicate that healthy aging up to age 75 does not significantly affect the shape of the HRF or the strength of BOLD neurovascular coupling. The majority of neuroimaging studies of cognitive aging use the cHRF for the convolution kernel [Grinband et al., 2008; Handwerker et al., 2004] making the implicit assumption that neurovascular coupling does not differ between groups and that any group differences in brain activity stem from neural sources. Our data supports this assumption for elderly subjects that have been extensively screened for clinical factors.

**AUTHOR CONTRIBUTIONS**

JG conceived and designed the study, analyzed the data, and drafted the manuscript; JG, JS, QRR, and YS collected the data and revised the manuscript.
REFERENCES


Grinband et al.


