Beat Fusion in the Brain of a Drummer: Expertise alters audiovisual brain activation

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Cerebral Cortex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>CerCor-2009-00857</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Original Articles</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>20-Oct-2009</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Pollick, Frank; University of Glasgow, Department of Psychology Petrini, Karin; University of Glasgow, Department of Psychology Dahl, Sofia; Aalborg University Copenhagen, Department of Media Technology McAleer, Phil; University of Glasgow, Department of Psychology McKay, Lawrie; University of Glasgow, Department of Psychology Rocchesso, Davide; IUAV University of Venice, Department of Art and Industrial Design Waadeland, Carl; Norwegian University of Science and Technology, Department of Music Avanzini, Federico; University of Padova, Department of Information Engineering Pupe, Aina; Indiana University, Department of Psychological and Brain Sciences</td>
</tr>
<tr>
<td>Keywords:</td>
<td>biological motion, multisensory, fMRI, music, audiovisual</td>
</tr>
</tbody>
</table>

Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.

drumming_movie.mov
Beat Fusion in the Brain of a Drummer: Expertise alters audiovisual brain activation

Frank E Pollick\textsuperscript{1}, Karin Petrini\textsuperscript{1}, Sofia Dahl\textsuperscript{2}, Phil McAleer\textsuperscript{1}, Lawrie S McKay\textsuperscript{1}, Davide Rocchesso\textsuperscript{3}, Carl Haakon Waadeland\textsuperscript{4}, Federico Avanzini\textsuperscript{5}, Aina Puce\textsuperscript{6,7}

\textsuperscript{1}Department of Psychology, University of Glasgow, United Kingdom,
\textsuperscript{2}Department of Media Technology, Aalborg University Copenhagen, Denmark,
\textsuperscript{3}Department of Art and Industrial Design, IUAV University of Venice, Italy,
\textsuperscript{4}Department of Music, Norwegian University of Science and Technology, Norway,
\textsuperscript{5}Department of Information Engineering, University of Padova, Italy,
\textsuperscript{6}Department of Radiology, West Virginia University School of Medicine, USA,
\textsuperscript{7}Department of Psychological and Brain Sciences, Indiana University, USA.

Address for correspondence:
Frank E Pollick
Department of Psychology
University of Glasgow
Glasgow, G12 8QB United Kingdom
Phone: +44 (0)141 330 3945
Fax: +44 (0)141 330 4606
Email: frank@psy.gla.ac.uk
We routinely and effortlessly merge audio and visual input to interpret the actions of others. In some instances, when mismatch occurs between the streams, additional processing may be required to interpret the stimulus. To study this issue, here we contrasted the ability of drummers and non-musical controls to detect mismatches in the synchrony or congruence of presented audiovisual information, using functional magnetic resonance imaging (fMRI) and a novel means of combining point-light visual displays with simulated audio of swing groove drumming. Large brain activation differences were seen between drummers and controls in the processing of audiovisual information. Synchrony detection engaged right temporal areas (MTG and STG) for drummers and a right frontal region (IFG) for controls. However, this same right IFG region was active in drummers, and not controls, when detecting incongruence, suggesting that only controls treat asynchrony detection as a cognitive task. In addition, for synchrony detection, drummers showed substantial task-induced deactivations during synchronous displays. Thus drumming expertise appears to provide drummers both with enhanced multisensory perceptual representations and management of processing resources.
Introduction

Interpreting the world around us requires combining multiple afferent streams of sensory information in a way that allows potentially disparate information about a common temporal event to be fused together. Psychophysical studies suggest that this fusion is driven by the salience of mid and high-level stimulus features, rather than being a purely data driven mechanism (2007). This is consistent with neuroimaging studies showing that familiarity can modulate how auditory and visual information are combined (Hein et al., 2007); suggesting that experience plays a vital role in multisensory perception and cognition. Indeed, the speech literature indicates that in one’s own native language, lip and facial movements of speech can augment speech interpretation (McGurk and MacDonald, 1976). Music, which shares many communicative and temporal rhythmic aspects with speech, is another domain where experience is important (Munte et al., 2002). For example, previous musical experience provides enhanced processing of auditory information at the level of the brainstem when vision is available (Musacchia et al., 2007), as well as enhancing activity at the cortical level in multisensory regions during the performance of an audiovisual task (Hodges et al., 2005).

The general network of brain areas involved in multisensory processing of audio and visual information has been revealed by studies comparing the processing of multisensory (audiovisual – AV) to unisensory (audio – A, or visual – V) stimuli. Despite different ways to interpret contrasts of brain activity during unisensory and multisensory processing (Holmes and Spence, 2005; Laurienti et al., 2005), converging evidence indicates that a distributed network of areas is involved (Calvert, 2001). This
network includes the superior temporal sulcus, the intraparietal parietal sulcus, regions of frontal cortex, insula and the claustrum.

The specific function of the different brain areas within this network has been studied by monitoring brain activity during the processing of mismatched versus matched AV stimulus pairs. AV pairings have included synchronous versus asynchronous displays which were compared on a wide variety of stimuli, including speech (Miller and D'Esposito, 2005; Ojanen et al., 2005) and flash-tone combinations (Bushara et al., 2001). Observed activation in a broad network of areas in the insula, posterior parietal cortex, prefrontal cortex and the superior temporal sulcus has been described in the form of increased activity during asynchronous versus synchronous presentations. Another type of pair-mismatching, comparing semantically congruent versus incongruent displays produced greater activation in the insula/claustrum as well as the anterior cingulate gyrus and adjacent medial prefrontal cortices (Laurienti et al., 2003; Naghavi et al., 2007). Additionally, a conjunction analysis of contrasts of bisensory versus unisensory presentations has revealed that the inferior frontal cortex is activated when attempting to match incongruent displays which are unfamiliar. Additional activity in the posterior superior temporal sulcus is seen when the displays are familiar (Hein et al., 2007).

Taken together, the results of the studies described above has led to the proposal that temporal regions are more involved for matching AV stimulation, possibly due to enhanced activation of stable AV object representations. In contrast, inferior frontal regions are more involved for the processing of incongruent stimuli, potentially reflecting increased cognitive control demands (Doehrmann and Naumer, 2008). These cognitive demands might reflect the nature of the task (i.e. a semantic or sensory mismatch),
however it has also been proposed that a switch of activity from posterior to frontal regions might be associated with more controlled processing when basic fusion mechanisms fail (Miller and D'Esposito, 2005). Here, we explored this issue by using carefully controlled audiovisual representations of drumming viewed by either drummers or non-musical controls (novices). The mode of visual stimulus presentation – point-light visual displays accompanied by synthetic sound generated by a physical model of impact – allowed control over both the synchrony and congruence between the visually presented action and its consequent sound.

Materials and Methods

Initial psychophysical study and subsequent fMRI experiments

Each participant completed one psychophysics experiment, adapted from previous research (Petrini, Dahl et al., 2009; Petrini, Russell et al., 2009), followed by two fMRI experiments. The purpose of the initial psychophysical study was to characterize the performance for each participant so that the stimuli used in the fMRI experiments could be optimized for that individual and also matched perceptually across all observers. The stimuli were point-light displays of drumming derived from motion-capture data of a drummer and were accompanied by sound generated from a physical model of impact. We describe the stimuli in more detail below.

Participants

All participants had either normal or corrected-to-normal vision. All participants had hearing levels that allowed them to clearly hear the auditory stimuli above the MRI
scanner noise. All participants volunteered for the study and were paid $25 for their participation. The study was approved by the Institutional Review Board of West Virginia University.

**Experts:** Eleven drummers (all males; age 35 ± 12 years) were studied. Seven of the 11 were presently either professional musicians and/or music teachers. All drummers were right handed as assessed by the Edinburgh Handedness questionnaire (Oldfield, 1971). The average number of years of drumming experience was 24 ± 11 years (range 13 to 45 years).

**Novices (control group):** We gender and age matched our novices relative to our expert group to obtain 11 males (age 35 ± 11 years) with no previous drumming experience and little or no previous musical training. Eight of the novices were right handed as assessed by the Edinburgh Handedness questionnaire (Oldfield, 1971).

**Stimulus creation**

The audiovisual stimuli were created in three distinct steps from initial motion capture data of a professional jazz drummer playing a swing groove beat (Petrini, Dahl *et al.*, 2009; Petrini, Russell *et al.*, 2009). Step 1 consisted of converting the 3D movement coordinates of the drummer into point-light displays of drumming actions using computer graphics. Step 2 consisted of converting the times and velocities of stick impact into a realistic stream of sound by using a naturalistic sound generation algorithm. Step 3 consisted of combining the visual point-light displays and the audio stream of drum beats by using video editing software.
Motion-capture data: Data were recorded from the movements of a professional jazz drummer playing a swing groove at 120 beats per minute with the accent on the second beat (Waadeland, 2006). Markers were placed at six locations including the tip of the drumstick, the level of the grip of the drumstick, and at the drummer’s shoulder, elbow, wrist and hand. The 3D coordinates of these marker locations were sampled at 240 Hz using a Proreflex 3D motion capture system.

Creation of visual display from motion capture data: The sampled 3D motion capture data were first downsampled to 60 Hz. They were then converted into a series of 2D images using Matlab and Psychtoolbox routines (Brainard, 1997; Pelli, 1997) running under OSX and utilizing OpenGL graphics with anti-aliasing enabled. White discs (luminance: 85 cd/m²; diameter: 2 mm) on a black background (luminance: 0.12 cd/m²) represented the drummer’s arm and drumstick (Figure 1). The drum head was represented using a thick white line (Figure 1), oriented 25 degrees from horizontal (width: 2.2 cm; height: 2 mm; luminance: 85 cd/m²). The image sequences were saved as video using an AVI file format at a 60Hz frame rate.

Creation of audio signal from motion capture data: The synthetic drumming sounds were obtained by an algorithm that took as input the times and velocities of a series of strikes and output the simulated audio signal (Fontana et al., 2004). To obtain a very natural sound the algorithm simulated the first 25 modes of a circular membrane. Both the time and impact velocity of a strike were derived by plotting the displacement and velocity of the drumstick tip marker against the time of the drummer performance and selecting, for
each impact, the frame at which the drumstick tip velocity changed from negative to positive ((Dahl, 2004); pp 765). To use only displacement and velocity perpendicular to the drumhead (Dahl, 2004; pp 765), data were rotated to a coordinate frame where horizontal was parallel to the drumhead and vertical perpendicular to the drumhead. These operations were performed on the 240 Hz, unfiltered displacements and velocities. The resulting sounds were saved as audio files (WAV format) with 15 second duration.

The audio files were created to either be congruent with the original movement data recordings or to be incongruent (see Figure 1). The congruent audio files were obtained by taking the impact times and impact velocities found in the original recorded movement and using these as input to the algorithm. For the incongruent audio files the identical impact times were used, but the set of velocities was randomly scrambled. Thus, for the congruent displays there was a natural covariation between the original movements of the drummer and the resulting sound (e.g. fast strikes – loud sounds, slow strikes – soft sounds); while for the incongruent displays the timing of the strikes was identical, but the natural covariation between the visually apparent velocity and the sound intensity was removed.

Combining audio and visual signals to create experimental stimuli: The 60Hz movies (AVI) and audio (WAV) files were imported into Adobe Premiere 1.5 where they were combined and synchronized to produce zero-lag audiovisual displays. Audiovisual displays with asynchronous audio and video were generated by either delaying the video with respect to audio by 4, 8, 12 and 16 frames (corresponding to lags: -66.67, -133.33, -200, and -266.67 ms) or by delaying audio with respect to video by 4, 8, 12 and 16
frames (corresponding to lags: 66.67, 133.33, 200, and 266.67 ms). This yielded a total of 9 audiovisual display files, including the one display with no delay.

To select the particular part of the audio and visual streams to use as stimuli, the initial 5 seconds of both files were first discarded to ensure the stabilization of the drummer performance. Next, a sound selection of 9 impacts was made, always starting from 2 frames before the first impact and ending at 1 frame before the tenth impact. The sound selection was kept constant for the 9 different lags, while the video was selected each time accordingly. The resulting audiovisual files each had a duration of 3 seconds. Here, it should be noted that since the offsets were relatively short compared to the entire duration available, there was always both audio and visual signal present at each instant of the 3 second displays. The resulting Quicktime movies were finally compressed by using Quicktime Pro 6. (An example of the original 0-delay movie is provided in the Supplementary Materials: drumming_movie.avi).

For the behavioral study, the audiovisual files were shown to participants by using Showtime (Watson and Hu, 1999), a component of the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) extension to Matlab. See also the work of Zhou and colleagues (Zhou et al., 2007) for Showtime used in multi-sensory integration on a Macintosh Powerbook G3 running OS9. For the fMRI study, Presentation V11 (Neurobehavioral Systems, CA) was used to present stimuli, to log behavioral responses and to record the times when MRI volumes were sampled with respect to stimulus delivery.

Experimental design
Psychophysical study: Here, we describe the two initial test sessions completed by each participant; these methods were adapted from closely related research (Petrini, Dahl et al., 2009; Petrini, Russell et al., 2009). Observers sat in a quiet, darkened room at a distance of approximately 100 centimeters from a computer monitor, and wore Beyer Dynamic DT Headphones. The maximum extent of the drumming figure was 5.6 cm in the horizontal direction, resulting in a visual angle of 3.2 degrees for the visual displays. The experiment consisted of 2 blocks of 90 stimuli run with a short rest in between. Each block contained 90 stimuli, consisting of a random ordering of the 2 conditions (congruent, incongruent) X 9 time lags (-266.67, -200, -133.33, -66.67, 0, 66.67, 133.33, 200, 266.67 ms) X 5 repetitions of each item. For the congruent condition drumbeats were generated using estimates of the striking velocity at the time of impact for that particular strike, while for the incongruent displays the sounds were generated using estimates of the striking velocity at the time of impact for a different strike.

After the experimenter had explained the task to the participant, the participant also read the task instructions on the computer screen. Participants were instructed to press “1” on the keypad if the drummer’s movements were perceived to be in synchrony with the sound, or press “3” if they were perceived as being asynchronous. After three training trials the experimenter left the participant alone to perform the experiment. After each movie, a short text message was displayed in the centre of the black screen to remind participants of the task and of the keys to use. When either response key was pressed the next movie was automatically displayed and so on until the end of the block.

The behavioral experiment produced 90 synchronization responses for both congruent and incongruent displays, which were distributed as 10 repetitions at each of
the 9 audiovisual timings – from sound preceding video by 266 ms to sound lagging video by 266 ms. The data were plotted as number of synchrony responses for each of the 9 timings and this data was fit with a Gaussian function. The peak of the Gaussian fit revealed the optimal timing for perceiving audiovisual synchrony for both congruent and incongruent stimuli for each participant, and the tail of the distribution furthest from the peak revealed the non-optimal timing. Typically, the optimal timing for perceiving the best audiovisual synchronization occurred when audio lagged video by around 90 ms, while the non-optimal timing for perceiving maximal desynchronization occurred when the audio led the video by 266 ms. Optimal and non-optimal videoclips were selected as stimuli in Experiments 1 and 2 of the fMRI study as appropriate (fMRI Experiment 1 used congruent optimal and congruent non-optimal stimuli, whereas fMRI Experiment 2 used congruent optimal and incongruent optimal stimuli). Hence, each participant was scanned with audiovisual stimuli that were \textit{a priori} individually detected as maximally synchronous or asynchronous.

\textbf{fMRI experiments:} A Windows PC (Dell Precision 690 PC) running Presentation V11.1 (Neurobehavioral Systems, CA) presented stimuli and logged behavioral responses and MRI data acquisition pulses for each acquired brain volume. The first MRI scanner pulse initiated the Presentation script, which presented an initial 8 second cue to the participants that the imaging run was beginning, during which a total of four MRI excitations were performed without MRI data acquisition to achieve steady-state magnetization. Visual clips were projected through the MRI scanner’s control room window onto a screen at the participant’s feet using a video projector (NEC Corporation, LT10 DLP). The participant lay supine in the MRI scanner and viewed the display
through a mirror mounted on the quadrature head coil. The stimulus size was adjusted so that the visual angle of the drummer was identical to the 3.2 degrees of visual arc used in the initial behavioral experiments. Audio was presented through a high quality sound card interface (CDX01, Digital Audio). A sound mixer (1642VLZ pro mixer, Mackie Inc.) and commercially available MR compatible electrostatic ear buds (STAX SRS-005 Earspeaker system; Stax LTD., Gardena, CA) worn under sound attenuating ear muffs were used to deliver sound to the participant’s ear.

In both experiments, a trial began with the 3 sec audiovisual clip and participants responded with a button press before movie offset. After video clip offset the trial ended with a blank (black) screen. On average, the blank screen was presented for 4 seconds, and this was varied randomly between 2 seconds to 6 seconds. A total of 50 trials were presented in randomized order during an imaging run which lasted 6 minutes and 40 seconds. Experiment 1 contained three runs and Experiment 2 contained two runs.

In fMRI Experiment 1, we varied the synchrony between sound and video to test the brain’s sensitivity to asynchrony for novices and drummers. There were 3 stimulus conditions: Synchronous (Synch), Asynchronous (Async) and a null condition containing just fixation. Participants indicated with a two button forced choice response whether sound and video were appropriately matched or mismatched. The synchronous and asynchronous items were taken from the congruent optimal and congruent non-optimal displays obtained in the psychophysical study.

In fMRI Experiment 2, we investigated differences in brain sensitivity to having the audio congruent or incongruent with the original striking kinematics. Participants were presented with 2 conditions at optimally synchronous timing: Congruent (Cong),
where the time-varying characteristics of the audio matched those of the recorded striking
kinematics, and Incongruent (Incong), where the original covariation between auditory
and visual time-varying characteristics was eliminated. Participants again indicated with
a two button forced choice response whether sound and video were appropriately
matched or mismatched. The incongruent item was taken from the optimal, incongruent
display obtained in the psychophysical study. The optimal, congruent display could have
been taken directly from the psychophysical experiment but we were concerned that since
this identical item had been used in Experiment 1 with the same participants, we would
potentially see reduced responses to it in Experiment 2. Thus, we created a new display
with the same optimal timing but taken from a non-overlapping section of the original
motion capture of the drumming performance.

**fMRI Data Acquisition**

**Functional Images:** We acquired blood oxygen level dependent (BOLD) activity
in a near whole brain acquisition using a gradient echo spiral in-out sequence (Glover and
Law, 2001) on a 3 Tesla Horizon HD MRI scanner (General Electric Medical Systems,
Inc.) using a quadrature headcoil. A total of 22 axial slices beginning from the vertex
(4mm thick with 1mm gap) were acquired with the following parameters: matrix = 128 X
128, FOV = 240mm, (in plane resolution = 1.875mm), bandwidth = 125, and TE/TR =
35/2000ms.

A total of 600 volumes were acquired for Experiment 1, whereas 400 volumes
were acquired for Experiment 2. In each participant, we acquired 200 volumes per run, so
that Experiments 1 and 2 had 3 and 2 imaging runs, respectively. Functional images used
in subsequent analyses consisted of averaged images of spiral-in and spiral-out
Expertise alters audiovisual brain activation trajectories, which were reconstructed off-line using routines written in C (courtesy of G. Glover, Stanford University, CA) and running under Linux (Fedora Core Release 5, Raleigh, NC). The averaged spiral in-out trajectories optimized sampling from brain regions prone to susceptibility artifacts and MR signal drop-out.

**Anatomical images:** In each participant, we acquired a near whole-brain T1-weighted anatomical volume with identical slice prescription to the functional images, and a whole-brain high-resolution anatomical SPGR volume (1.5mm x 0.9375mm x 0.9375mm, FOV = 240, matrix 256 X 256, 124 slices).

**fMRI Data Analysis**

**Data pre-processing:** Functional and anatomical images were analyzed using Brain Voyager QX 1.10 (Brain Innovation, Maastricht, The Netherlands). Functional imaging data were pre-processed by performing a slice scan time correction (using sinc interpolation), linear trend removal and temporal high-pass filtering to remove low-frequency non-linear drifts of 3 or fewer cycles per time course. In addition, 3D motion correction (6df) was performed to detect and correct for small head movements and estimated translation and rotation were inspected and translation never exceeded 3mm. Functional slices were coregistered to the anatomical volume using manual alignment to obtain optimal fit and the resulting functional and anatomic data were transformed into Talairach space.

**First level analysis:** Analyses were performed on the data of individual subjects using multiple linear regression of the BOLD-response time course in each voxel using three predictors (Synchrony, Asynchrony and Blank) in Experiment 1 and two predictors (Congruent, Incongruent) in Experiment 2. Predictors’ time courses were adjusted for the
hemodynamic response delay by convolution with a hemodynamic response function (Boyton et al., 1996).

To examine the time-course of BOLD activity associated with the different stimulus conditions we performed event-related averaging for the different experimental conditions. This resulted in a plot of percentage BOLD signal change relative to a baseline value. The baseline value was calculated by first finding the pre-period BOLD value at each trial and then averaging all these pre-period values over the entire time course.

Group analysis: Fixed-effect analyses were performed on each groups’ data (n=11 for each group). Multisubject statistical maps were thresholded using the false discovery rate (FDR; Genovese et al, 2002). For ROI-based analyses, t statistics and corresponding p values are reported.

Results

In both experiments, the same individuals from two groups of novices and experienced drummers were scanned while providing trial-by-trial judgments of whether the sight matched the sound of a point-light drumming display. In Experiment 1 the stimuli varied in the synchrony of the audiovisual signal, and in Experiment 2 they varied in the congruence of the audiovisual signal, as depicted in the schematic shown in Figure 1.

Experiment 1: Effects of Audiovisual Synchrony

In Experiment 1 participants were presented with an audiovisual swing-groove drumming display that was either synchronous or asynchronous, with a three second
Expertise alters audiovisual brain activation

duration, and asked to press a response button to indicate whether the sound matched or
mismatched the sight. Trials with mismatches and matches were presented in a random
sequence while participants lay supine in the MRI scanner. The timings of these displays
were determined immediately prior to entering the MRI scanner by a separate
psychophysical experiment where optimal timings to produce synchronous and
asynchronous responses were measured for each individual participant using previously
developed methods (Petrini, Dahl et al., 2009). The results of participants’ behavioral
judgments during MRI scanning indicate that both the novices and the experienced
drummers were performing above 90% correct (Figure 2a). Here, correct is defined
perceptually rather than physically, i.e., reporting the predetermined perceptually
synchronous and asynchronous displays as synchronous and asynchronous respectively.
A comparison of the performance between novices and drummers failed to show a
significant difference between the two groups’ abilities to discriminate the synchronous
display as matching and the asynchronous one as mismatching (t(20)=0.167; p=0.869).
This result is important since it indicates that differences in brain activity between the
groups cannot be attributed to the novice group experiencing the task as being more
difficult than the drummers.

------- Figure 2 -------

In the fMRI data, we observed brain regions where activation was greater for
asynchronous than synchronous displays (Asynch>Synch) for both drummer and novice
groups (Figure 2b). The anatomical location and details of the activated foci are listed in
Table 1. Likewise, we observed greater activation when we contrasted synchronous
versus asynchronous displays (Synch>Asynch), which are depicted in Figure 2c and Table 2. Subsequent visual examination of the timecourse of activity for each of these regions indicated that the significant difference in activation was produced by clear positive or negative BOLD responses. Negative BOLD has been interpreted as a task-induced deactivation (TID) which has often been reported as a characteristic of the default network (Raichle et al., 2001). The TID has been postulated to occur as a response of the default network deactivating during the processing of external stimuli (Binder et al., 1999; McKiernan et al., 2003; Shulman et al., 1997). In Tables 1 and 2 we include a column to indicate whether the differential activity could be accounted for by TID. In Figure 3a we have plotted drummers’ timecourses of activation for Experiment 1 in response to the blank, synchronous and asynchronous displays for two regions of positive activation, and in Figure 3b for two regions that exhibited TID. (Plots of the timecourses of other regions showing positive activation and TID can be viewed in the Supplementary Figures 1 and 2, respectively.) This tendency towards TID occurred in both the drummers and the novices, albeit with notably different patterns. For drummers, TID for the synchronous displays in the Asynch>Synch contrast was found in the right hemisphere for the lingual gyrus, MTG (BA 39), precuneus and posterior cingulate, and in the left hemisphere for the anterior cingulate, SFG and MOG. For novices, a TID for the asynchronous displays in the Synch>Asynch contrast was found in the posterior cingulate of the right hemisphere and the SFG of the left hemisphere.

-------- Figure 3 --------
Both drummers and novices activate an extensive network of regions when processing synchrony that show negligible overlap between the two groups. For the contrast of Asynch>Synch, drummers show activity in both hemispheres and relatively small activations in frontal areas compared to large areas of activation in more posterior regions known to be involved in multisensory processing such as the STG and MTG. In contrast, novices show a response lateralized largely to the left hemisphere and the frontal lobe with an additional large activation in the right IFG. For the Synch>Asynch comparison, there is again a clearly different activation pattern for the drummers and novices. The drummers show only activation in the left IFG while the novices show activation in subcortical regions along with activation in the posterior cingulate and SFG associated with a TID response for asynchronous displays.

Effects of Audiovisual Congruence – Experiment 2

In Experiment 2 participants were presented with 3 second videos of an audiovisual swing-groove drumming display that was either congruent or incongruent, and were asked to press a response button to indicate if the sound matched or mismatched the sight during the fMRI scan. Similar to the first experiment, the occurrence of trial type was randomized over the course of the experiment. For each individual participant, the timings of these displays were determined immediately prior to entering the scanner by a behavioral experiment so that each individual would perceive both the congruent and incongruent displays as maximally synchronous. Thus, the fundamental difference between displays was that the congruent displays had the original mapping of strike velocity to physically-modeled sound, while the incongruent displays had a random mapping between strike velocity and physically modeled sound. The results of
participants’ judgments of whether or not the sight and sound matched are presented in Figure 4a, where a significant difference can be seen between the drummers and novices (t(20)=3.902; p=0.001), with the novices not being significantly different from chance (t(10)=0.111, mean = 50.54, Standard Error=4.90). This result would suggest the possibility that no difference in brain activation would be found for the novices across the experimental conditions.

Indeed, for the novices there was no significant brain activation when incongruent and congruent displays (Incong>Cong) were contrasted. A similar result was observed when the contrast was reversed (Cong>Incong). However, drummers did exhibit activation for the contrasts of Incong>Cong and Cong>Incong. The activation foci are shown in Figure 4b and their anatomical locations and co-ordinate maxima in standard space are provided in Table 3. Examination of the timecourse of activation for the congruent and incongruent trials for each of these regions again indicated that the significant difference in activation could be attributed to either a positive BOLD response or TID. However, comparison with a blank condition was not available to confirm this supposition. For the contrast of Incong>Cong in Experiment 2, drummers showed only positive BOLD activations, which is similar to the pattern shown by the novices in Experiment 1 for the Asynch>Synch comparison. Further examination of the data of the drummers in Experiment 2 suggested a TID for incongruent displays in the Cong>Incong
contrast in the right hemisphere for the posterior cingulate and MFG and in the left hemisphere for the MTG.

Drummers activated an extended network of brain regions known to be involved in the processing of congruence that included frontal, parietal and temporal cortex. With a single exception, these regions did not match those found for either the drummer or novice group in Experiment 1. The one exception is the right IFG which was also found for the novices in the Asynch>Synch contrast.

Discussion

Our results demonstrate an important role for experience in shaping brain responses and behavior to multisensory stimuli. While both drummers and novices show activation in brain regions often associated with multisensory processing, their patterns of activation were distinctly different. Drummers exhibited a more extensive network of activations, a result consistent with behavioral data which demonstrate that, for combining the sight and sound of drumming actions, drummers are more narrowly tuned and stable across variations in tempo (Petrini, Dahl et al., 2009), and can robustly fuse audio and visual information even when visual conditions are substantially degraded (Petrini, Russell et al., 2009). The neural basis of these perceptual advantages would seem not to arise from a single mechanism, but rather an array of processing differences in different brain regions. The following paragraphs describe how expertise influences audiovisual processing in these different brain regions.
The rIFG, an area known for its involvement in semantic processing of audiovisual stimuli (Doehrmann and Naumer, 2008), showed clear differences between groups in its response to the tasks of judging synchrony and congruence. Novices showed activity in this area for the contrast of asynchrony versus synchrony, while drummers showed activity in the identical area for the contrast of incongruent versus congruent. To understand this difference between groups it is useful to consider the relative semantic content of the displays (Hein et al., 2007; Noesselt et al., 2007; Van Atteveldt et al., 2007). Both asynchrony and incongruence were created by differences in the physical match between the audio and visual streams. However, the asynchronous displays were grossly mismatching in timing of physical signals while the incongruent displays were matched for optimal timing but had a subtle mismatch in the cause and effect of striking velocity and resulting beat accent. Thus, on a continuum of semantic difficulty the asynchronous displays posed the simple problem of whether sound and sight matched at all, while the incongruent displays posed the sophisticated problem of whether the fine structure of the sight and sound were matched. With this in mind we can consider the suggestion of Doehrmann and Naumer (2008) that the matching of familiar, possibly overlearned, AV stimuli is performed in a network consisting of lateral temporal and inferior frontal regions with a higher sensitivity of the lateral temporal areas to semantically matching combinations and of the inferior frontal areas to mismatching combinations. Our results are thus consistent with the view that the predominantly frontal activation for the novices implies that because drumming is unfamiliar, detecting asynchrony engages these semantic mechanisms of mismatch detection. In contrast, for the drummers the detection of asynchrony can be accomplished by more perceptual
mechanisms, with frontal mechanisms becoming involved for cases such as congruence when the mismatch involves a semantic level of processing.

The MTG and STG were found to be active only in drummers for the contrast of asynchronous versus synchronous displays and in the left STG for the contrast of incongruent versus congruent displays. These results for asynchronous versus synchronous displays agree with data obtained in musical conductors when performing temporal order judgments (Hodges et al., 2005). Hodges and colleagues found that the conductors had extensive activations in the temporo-parietal regions of cortex that were not activated in their non-musical control subjects. These differences were obtained using relatively simple perceptual stimuli that consisted of visual circles and beeps. The activation similarities in our study extend these differences in MTG and STG to more complex musical stimuli. The left-lateralized response found in left STG for incongruent versus congruent displays is similar to that found in jazz musicians when presented with rhythmic violations (Vuust et al., 2005). Vuust and colleagues used MEG to explore mismatch negativity and found a left-lateralized mismatch negativity which was localized to the posterior regions of the temporal cortex. Their interpretation was that expert jazz musicians use violations of rhythm to communicate during improvisational performance, and thus similar to language, a left lateralized mismatch negativity is obtained. Such an explanation can be adapted to our incongruent displays if we assume our drummers were seeking meaning in the presented audiovisual displays. Consistent with this are the spontaneous comments of one drummer who on debriefing reported musically sophisticated interpretations of the incongruent displays.
It is known that musical training impacts the encoding of audiovisual information in subcortical structures (Musacchia et al., 2007) and that activation in the insula has also been interpreted to occur as a consequence of its connection with brainstem regions that support subcortical detection of audiovisual correspondence (Bushara et al., 2001). These reports are relevant to interpreting the results of the contrast of synchronous versus asynchronous displays in Experiment 1 that yielded a single cortical activation (left IFG) for drummers and 3 subcortical activations in novices. While activation in a more caudal region of left IFG has been reported as a consequence of experiencing a recently learned musical piece (Lahav et al., 2007), our left IFG activation extended into the insula and bordered the coordinates obtained by Bushara and colleagues for an audiovisual integration task (Bushara et al., 2001). Thus, our findings taken together with those of Musacchia et al. (2007) and Bushara et al. (2001) would suggest that drumming expertise changes corticofugal influences on brainstem activity as well as producing more efficient encoding in the brainstem.

Several areas reported in Experiment 1 for the drummers’ contrast of asynchrony versus synchrony can be related to the putative default network in the brain (Buckner et al., 2008; Fox et al., 2005; Raichle et al., 2001). Consistent with this is that these areas resulted from task-induced deactivations (i.e. the BOLD response for the synchronous displays was negative while the asynchronous condition remained around baseline), a fact that was confirmed by comparison to activity levels during the blank condition where BOLD signal levels were positive. If one considers only the timecourses for the blank and synchronous conditions, an explanation consistent with the default network is possible; namely, that these areas are active when they are not stimulus driven and
become ‘deactivated’ when sensory processing occurs. However, such an explanation
does not explain why the timecourse of BOLD activation for the asynchronous displays
does not also go negative, nor why novices do not show the same result. A possible
explanation of this might be that musicians can rapidly focus on processing musical
stimuli, thereby definitively leaving a default mode of processing. This hypothesis of
increased focus by musicians has also been advanced to explain task induced
deactivations in the default network during musical performance (Limb and Braun, 2008;
Parsons et al., 2005). What distinguishes our results from these other studies is that in
our study similar deactivations were obtained during a task that did not require any
production of music. This suggests that musical expertise likely influences the activity of
the default network not only for music production, but also the multisensory perception,
and potentially interpretation, of musical performance. Whether such a result would be
observed if only auditory or visual presentation was experienced by these expert subjects
remains an open question.

In summary, our results provide several insights into how musical expertise might
be expressed in the brains of drummers when experiencing an audiovisual drumming
performance by another performer. Foremost, drummers appear to recruit more
semantically driven brain areas such as the right IFG only when subtle discriminations
are required, whereas novices rely on such mechanisms for even basic synchronicity
judgments. It may be that there is more extensive multisensory processing in the
temporal cortex of expert drummers which is enhanced by an ability to clearly “switch
off” the default network when presented with synchronous audiovisual presentations.
There is also evidence to suggest that these expert subjects are more efficient at
subcortical processing of the audiovisual signal. The current results, however, do not allow us to generalize to other musicians, as it has been suggested that differences in neurocognitive profile might reflect unique musical specialization (Tervaniemi, 2009). Nor do our data indicate whether these functional differences arise from anatomical differences that themselves arise from brain plasticity (Munte et al., 2002). Our data do provide several explanations as to why drummers are robust in their abilities to fuse the sight and sound of drumming (Petrini, Dahl et al., 2009; Petrini, Russell et al., 2009) and indicate that drumming is a potentially rich area in which to study multisensory processing and expertise.

Acknowledgements

We would like to acknowledge the support of the British Academy (LRG-42455) and the ESRC (RES-060-25-0010). We thank all the drummers for their assistance and in particular Robert Davis II and Paul Evans for their helpful discussions. We thank Mary Pettit for administrative assistance.
References


Table 1  
Experiment 1: Brain regions activated in the Asynchrony > Synchrony\(^a\) contrast for both drummers and novices.

<table>
<thead>
<tr>
<th>Anatomical region</th>
<th>Hemisphere</th>
<th>Talairach Co-ordinate</th>
<th>Number of voxels</th>
<th>Effect size(^b)</th>
<th>T</th>
<th>P</th>
<th>TID(^c)</th>
<th>BA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Drummers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STG</td>
<td>Right</td>
<td>59, -52, 13</td>
<td>6385</td>
<td>3.88</td>
<td>0.0003</td>
<td>N</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Lingual Gyrus</td>
<td>Right</td>
<td>8, -80, 0</td>
<td>18499</td>
<td>3.66</td>
<td>0.0004</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTG</td>
<td>Right</td>
<td>48, -65, 14</td>
<td>673</td>
<td>3.60</td>
<td>0.0005</td>
<td>Y</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>MFG</td>
<td>Right</td>
<td>37, -9, 42</td>
<td>1282</td>
<td>3.79</td>
<td>0.0004</td>
<td>N</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>MFG</td>
<td>Right</td>
<td>22, 34, 42</td>
<td>1896</td>
<td>3.55</td>
<td>0.0006</td>
<td>N</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Precuneus</td>
<td>Right</td>
<td>8, -55, 39</td>
<td>3044</td>
<td>3.64</td>
<td>0.0004</td>
<td>Y</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Posterior Cingulate</td>
<td>Right</td>
<td>13, -54, 17</td>
<td>733</td>
<td>3.49</td>
<td>0.0006</td>
<td>Y</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>MFG</td>
<td>Right</td>
<td>12, 50, 14</td>
<td>436</td>
<td>3.60</td>
<td>0.0005</td>
<td>N</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Anterior cingulate</td>
<td>Left</td>
<td>-7, 44, 6</td>
<td>1288</td>
<td>3.56</td>
<td>0.0006</td>
<td>Y</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>SFG</td>
<td>Left</td>
<td>-20, 24, 44</td>
<td>992</td>
<td>3.76</td>
<td>0.0004</td>
<td>Y</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>MOG</td>
<td>Left</td>
<td>-33, -77, 22</td>
<td>750</td>
<td>3.46</td>
<td>0.0006</td>
<td>Y</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Precentral Gyrus</td>
<td>Left</td>
<td>-32, -25, 65</td>
<td>452</td>
<td>3.73</td>
<td>0.0004</td>
<td>N</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>MTG</td>
<td>Left</td>
<td>-54, -47, 6</td>
<td>988</td>
<td>3.53</td>
<td>0.0005</td>
<td>N</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td><strong>Novices</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IFG</td>
<td>Right</td>
<td>45, 18, 15</td>
<td>14986</td>
<td>4.11</td>
<td>0.0002</td>
<td>N</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>MFG</td>
<td>Left</td>
<td>0, 16, 47</td>
<td>3556</td>
<td>3.82</td>
<td>0.0003</td>
<td>N</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Precentral Gyrus</td>
<td>Left</td>
<td>-33, 0, 32</td>
<td>2225</td>
<td>3.85</td>
<td>0.0003</td>
<td>N</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Claustrum</td>
<td>Left</td>
<td>-32, 17, 2</td>
<td>2845</td>
<td>3.99</td>
<td>0.0002</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IFG</td>
<td>Left</td>
<td>-52, 37, 9</td>
<td>563</td>
<td>3.58</td>
<td>0.0004</td>
<td>N</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>IFG</td>
<td>Left</td>
<td>-57, 12, 16</td>
<td>677</td>
<td>3.59</td>
<td>0.0004</td>
<td>N</td>
<td>44</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Statistical test used for ROI selection (maps thresholded at q(FDR)< 0.05).

\(^b\) Effect size = average t-value for all voxels in the ROI.

\(^c\) Task Induced Deactivation (TID): Y=yes; N=no.
Table 2
Experiment 1: Brain regions activated in the Synchrony > Asynchrony\(^a\) contrast for drummers and novices.

<table>
<thead>
<tr>
<th>Anatomical region</th>
<th>Hemisphere</th>
<th>Talairach Co-ordinate</th>
<th>Number of voxels</th>
<th>Effect size(^b)</th>
<th>TID(^c)</th>
<th>BA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Drummers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IFG</td>
<td>Left</td>
<td>-32, 17, -6</td>
<td>1144</td>
<td>3.84</td>
<td>0.0003</td>
<td>N</td>
</tr>
<tr>
<td><strong>Novices</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Culmen</td>
<td>Right</td>
<td>27, -50, -23</td>
<td>802</td>
<td>3.62</td>
<td>0.0004</td>
<td>N</td>
</tr>
<tr>
<td>Lentiform Nucleus</td>
<td>Right</td>
<td>31, -8, -2</td>
<td>926</td>
<td>3.59</td>
<td>0.0004</td>
<td>N</td>
</tr>
<tr>
<td>Posterior Cingulate</td>
<td>Right</td>
<td>5, -52, 18</td>
<td>9694</td>
<td>3.81</td>
<td>0.0003</td>
<td>Y</td>
</tr>
<tr>
<td>SFG</td>
<td>Left</td>
<td>-12, 37, 49</td>
<td>719</td>
<td>3.64</td>
<td>0.0004</td>
<td>Y</td>
</tr>
<tr>
<td>Lentiform Nucleus</td>
<td>Left</td>
<td>-24, -3, -3</td>
<td>429</td>
<td>3.63</td>
<td>0.0004</td>
<td>N</td>
</tr>
</tbody>
</table>

\(^a\) Statistical test used for ROI selection (maps thresholded at q(FDR)< 0.05).

\(^b\) Effect size = average t-value for all voxels in the ROI.

\(^c\) Task Induced Deactivation (TID): Y=yes; N=no.
Table 3
Experiment 2: Activated brain regions for drummer subjects for contrasts across incongruous and congruous stimulus conditions.

<table>
<thead>
<tr>
<th>ROI</th>
<th>Hemisphere</th>
<th>Talairach Co-ordinate</th>
<th>Number of voxels</th>
<th>Effect size&lt;sup&gt;b&lt;/sup&gt;</th>
<th>TID&lt;sup&gt;c&lt;/sup&gt;</th>
<th>BA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Incong &gt; Cong</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPL</td>
<td>Right</td>
<td>53, -47, 38</td>
<td>3644</td>
<td>3.56</td>
<td>0.0007</td>
<td>N</td>
</tr>
<tr>
<td>MTG</td>
<td>Right</td>
<td>60, -43, -10</td>
<td>1549</td>
<td>3.67</td>
<td>0.0005</td>
<td>N</td>
</tr>
<tr>
<td>IFG</td>
<td>Right</td>
<td>47, 16, 18</td>
<td>24183</td>
<td>4.07</td>
<td>0.0003</td>
<td>N</td>
</tr>
<tr>
<td>Thalamus</td>
<td>Right</td>
<td>3, -24, 1</td>
<td>1505</td>
<td>3.44</td>
<td>0.0008</td>
<td>N</td>
</tr>
<tr>
<td>SFG</td>
<td>Right</td>
<td>1, 10, 54</td>
<td>7709</td>
<td>4.40</td>
<td>0.0002</td>
<td>N</td>
</tr>
<tr>
<td>Insula</td>
<td>Left</td>
<td>-39, 13, 1</td>
<td>13296</td>
<td>3.91</td>
<td>0.0004</td>
<td>N</td>
</tr>
<tr>
<td>Precuneus</td>
<td>Left</td>
<td>-31, -61, 40</td>
<td>1856</td>
<td>3.52</td>
<td>0.0007</td>
<td>N</td>
</tr>
<tr>
<td>STG</td>
<td>Left</td>
<td>-52, -45, 20</td>
<td>2376</td>
<td>3.53</td>
<td>0.0007</td>
<td>N</td>
</tr>
<tr>
<td>IPL</td>
<td>Left</td>
<td>-48, -41, 37</td>
<td>414</td>
<td>3.35</td>
<td>0.0009</td>
<td>N</td>
</tr>
<tr>
<td><strong>Cong &gt; Incong</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postcentral Gyrus</td>
<td>Right</td>
<td>38, -24, 45</td>
<td>453</td>
<td>3.42</td>
<td>0.0008</td>
<td>N</td>
</tr>
<tr>
<td>Posterior Cingulate</td>
<td>Right</td>
<td>2, -58, 18</td>
<td>4028</td>
<td>3.64</td>
<td>0.0006</td>
<td>Y</td>
</tr>
<tr>
<td>MFG</td>
<td>Right</td>
<td>1, 50, 15</td>
<td>666</td>
<td>3.59</td>
<td>0.0006</td>
<td>Y</td>
</tr>
<tr>
<td>MTG</td>
<td>Left</td>
<td>-45, -76, 22</td>
<td>1395</td>
<td>3.55</td>
<td>0.0007</td>
<td>Y</td>
</tr>
</tbody>
</table>

<sup>a</sup> Statistical test used for ROI selection (maps thresholded at q(FDR)< 0.05).

<sup>b</sup> Effect size = average t-value for all voxels in the ROI.

<sup>c</sup> Task Induced Deactivation (TID): Y=yes; N=no.
Figure Legends

Figure 1. Experiments 1 and 2 stimulus condition schematic. In the top center of the figure a frame from the point-light display is presented. The point-light dots represent the drummer’s arm beginning at the shoulder joint. Note that the white line outlining the drummer is presented here for clarity only and did not appear in the presented stimulus. The left column displays the attributes of the stimuli used in Experiment 1. The displays had an audio signal congruent with the visual signal and were presented either in a perceptually optimal synchronous or asynchronous manner. The attributes of the visual motion, in terms of the relationship of the original motion velocity relative to implied velocity, appear on the top plot, and the sound waveforms appear directly under that. The lowermost panel of waveforms depicts the relationship of the timing of the auditory and visual stimuli relative to one another. The right column depicts stimulus attributes in Experiment 2, where the audio and visual signals were either congruent or incongruent and were always presented in a perceptually optimal synchronous manner. The top plots show the relationship of the original motion velocity relative to implied velocity – with the congruent and incongruent condition being displayed on the left and right, respectively. The sound waveforms and timing of the audio stimuli relative to the video appear as the last two items in the column.

Figure 2. Experiment 1: Group behavioral and activation data. A) Response accuracies (% correct responses) for data collected in the MRI scanner are shown for novices and drummers in the form of detection of audio and visual signal matches and mismatches. In this experiment all displays were congruent and were either optimally synchronous or asynchronous. The dashed line at 50% indicates chance performance.
B) Group activation data depicting differentially activated voxels for the Asynchrony > Synchrony contrast. Axial slices show activation foci at six Talairach z co-ordinates. Significant activations for the controls and the drummers are shown in blue and yellow, respectively. C) Differentially activated voxels for the contrast Synchrony > Asynchrony are shown in selected axial slices at six Talairach z co-ordinates.

Figure 3. Experiment 1: Drummers’ timecourses of activation for various regions active for Asynchrony > Synchrony. Three curves of percentage BOLD signal change in each plot as a function of time across the three stimulus conditions: blank (dotted line), synchronous (broken line) and asynchronous (continuous line). A) Two examples of positive activations in the left (L) MTG and left precentral gyrus. The BOLD response is larger for the asynchronous relative to the synchronous condition, whereas the blank condition BOLD response becomes negative. B) Two examples of task-induced deactivation (TID) in the right (R) lingual gyrus and the right precuneus. TID is indicated by the synchronous curve (broken line) dipping negative, the asynchronous condition staying approximately at baseline and the blank condition curve peaking in a positive direction. Additional timecourses for additional brain regions are shown in Supplementary Figures 1 and 2.

Figure 4. Experiment 2: Group behavioral and activation data for drummer and control subjects. A) Response accuracy (% correct responses) for the task performed (detection of audio and visual signal matches and mismatches) in the scanner by the controls and drummers. The presented audiovisual displays were all optimally synchronous and either congruent or incongruent. The dashed line represents chance performance. B) Activation data for the Incongruent > Congruent (yellow) and
Congruent > Incongruent (purple) contrasts are shown for the drummers. No significantly activated voxels were identified in controls for these contrasts.
Figure 1.
Figure 2.

A

![Bar chart showing percentage of correct responses for Controls and Drummers.]

B

Asynch > Synch

![Images of brain scans at different Z levels showing activity changes.]

C

Synch > Asynch

![Images of brain scans at different Z levels showing activity changes.]

Figure 3.

--- Blank  --- Synchronous  --- Asynchronous

**LMTG**

(a)

**LPCG**

Time (Seconds)

% BOLD signals

**RLG**

Time (Seconds)

% BOLD signals

**R Precuneus**

Time (Seconds)
Figure 4.
Supplementary Material: Positive Activation

Figure S1: Experiment 1 drummers’ timecourse plots of activation for additional regions active for the contrast of Asynchrony > Synchrony (positive BOLD for Asynchronous). Three curves of percentage BOLD signal change in each plot correspond to the three stimulus conditions: blank, synchronous and asynchronous. These figures show the cases when activation due to the asynchronous condition resulted in a positive BOLD response that was greater for the asynchronous than the synchronous displays.
**Supplementary Material: Task-Induced Deactivation (TID)**

--- Blank --- Synchronous — Asynchronous

Figure S2: Experiment 1 drummers’ timecourse plots of activation for additional regions active for the contrast of Asynchrony > Synchrony (negative BOLD for Synchronous). Three curves of percentage BOLD signal change in each plot correspond to the three stimulus conditions: blank, synchronous and asynchronous. These figures show the cases when activation due to the asynchronous condition resulted in a negligible BOLD response while the synchronous condition produced a negative response. The blank condition produced a positive response. These conditions are consistent with Task Induced Deactivation (TID).