Neural Correlates of Joint Drumming : An fMRI study

September 16, 2011



Theresa Marie Veltri

The University of Sheffield

Department of Psychology

SUPERVISOR: Lawrence Parsons

Dissertation

Submitted in Partial Fulfillment of the Requirements for the Degree of

Master's of Cognitive and Computational Neuroscience

Acknowledgements

I would like to thank Renee Timmers for the advice and support that she has provided unconditionally. Renee, I thank you most for your patience and for your willingness to comprise when it comes to our differences in communication style. I also thank you for your ability to be flexible with my ever-changing moods. And yes, I also continue to thank you for your sense of humor, which never ceases to surprise me. How can someone so calm and cool-headed contain such wit and cheeky facial expressions? Yes indeed, it has truly been a pleasure.

I would also like to thank my parents for their comic relief. Thank you again for listening to me complain and for pretending to understand what "the cognitive neuroscience of music" actually is. Without your emotional and financial support this degree, more so than the last, would have never been possible.

I would like to also thank Josias Bermea for your patience and understanding. Getting a post-graduate degree is difficult and time consuming. Doing so overseas adds all the more burden. Thank you for being understanding when it seemed I was avoiding you for days and most of all I thank you most for letting me be myself.

I would like to thank Bhavin Parekh, whose encouragement motivated me to finish this degree. Thank you for teaching me what science and research really are and for teaching me how to ask the big questions in music psychology.

I thank Ying Zheng for her help with SPM. I definitely could not have finished this degree without your help.

I also thank Kevin Gurney John Porrill for their patience, which I know I tested on a number of occasions.

I also thank Larry Parsons for his comic relief. During this degree I often missed the insulting humor that Americans are so used. It was nice to exchange insults with you over a cup of Starbucks.

I would also like to thank Deepti Parmar and Nellely Cervantes for keeping my head sane and my stomach full.

Lastly, I would like to thank my past mentors for their encouragement throughout my education: Bryan Buffaloe, Andrea Yun, and Stacey Davis. Your kindness and support have not been forgotten.

Abstract

Music performance is commonly thought of as a social activity. The coordination necessary for music performance amongst two or more individuals may have evolved from the need to emotionally communicate, and therefore rhythmically synchronize, with others. Because music is a social activity research investigating individual's sensorimotor synchronization abilities does not offer a complete understanding of how musicians coordinate during a joint performance. Therefore, the current study aims to examine joint drumming performances in 8 musicians though a behavioral and fMRI study. Behavioral findings show that isochronous drumming results in the least amount of synchronization variability, while the fMRI findings were inconclusive. Further research may be interested in examining the brain and behavioral effects of joint drumming using more sophisticated audio recordings and different brain imaging techniques.

Contents

1 Behavioral Research on Entrainment **1**

1.1 Entrainment . . . 1

1.2 Rhythmic Coordination . . . 1

1.3 Sensorimotor Synchronization . . . 2

1.4 Duet drumming . . . 3

2 The Neural correlates of joint rhythm and action **4**

2.1 Evolution . . . 4

2.2 Timing and movement . . . 4

2.3 Rhythmic complexity . . . 4

2.4 Neural correlates of synchronization . . . 5

2.5 Neural correlates of practice . . . 6

2.6 Neural correlates of joint action . . . 6

2.7 Current study . . . 7

3 Techniques **9**

3.1 Use of fMRI . . . 9

3.2 Physiology underlying the BOLD signal . . . 10

3.3 Sparse sampling technique . . . 10

3.4 Analysis technique: GLM . . . 12

4 Method **14**

4.1 Participants . . . 14

4.2 Drumming Stimuli . . . 14

4.3 Materials . . . 15

4.4 Experimental tasks . . . 15

4.5 Procedure . . . 16

4.6 fMRI scanning . . . 16

4.7 Data Analysis . . . 17

4.8 Preprocessing . . . 18

4.9 First level analysis . . . 18

4.10 Second level analysis . . . 18

5 Results **19**

5.1 Behavioural Data . . . 19

5.1.1 Level of Synchronization in Complexity and Joint Action Conditions 19

5.1.2 Levels of Synchronization for Isochronous, Double, and LSS conditions . . . 20

5.2 Brain Imaging Data . . . 22

5.2.1 Simple Verse Complex Condition . . . 22

5.2.2 Overall Activation of Joint Action . . . 22

5.2.3 Neural Correlates of the Human-Human Joint Action Condition and Solo Condition . . . 24
5.2.4 Neural Correlates of the Human-Human Condition and Human-Computer Condition . . . 24
5.2.5 Neural Correlates of the Solo and Human-Computer Conditions . 24

6 Discussion **25**

6.1 Aim . . . 25
6.2 Behavioral Data . . . 25
6.2.1 Level of Synchronization in Complexity and Joint Action Conditions 25
6.2.2 Levels of Synchronization for Isochronous, Double, and LSS conditions . . . 25
6.3 Brain Imaging Data . . . 26
6.3.1 Neural Correlates of the Simple Condition and Complex Condition 26
6.3.2 Neural Correlates of the Human-Human Joint Action Condition and Solo Condition . . . 26
6.3.3 Neural Correlates of the Human-Human Condition and Human-Computer Condition . . . 27
6.3.4 Neural Correlates of the Solo and Human-Computer Conditions . 27
6.4 Limitations . . . 27
6.5 Future Research . . . 28
6.6 Conclusion . . . 29

7 Bibliography **31**

1 Behavioral Research on Entrainment

1.1 Entrainment

We have a great ability to coordinate our actions with others. Remarkably, we can coordinate quickly with little or no conscious effort (Konvalinka, Vuust, Roepstorff & Firth, 2007, 2010), even in the absence of explicit instruction (Oullier et al., 2008; Richardson, Marsh, & Schmidt, 2005, Schmidt & O'Brien, 1997). For example, partners synchronize body movements when walking side by side and during conversations. Humans also synchronize their movements with the rhythms of external sounds, such as a beating drum (Wallin, Merker, & Brown, 2000), or sounds produced by others, especially when dancing (Nettl, 2005), or more relevantly, when performing in a musical ensemble (Newman-Norlund, Noordzij, Meulenbroak, & Bekkering, 2007a). Humans' ability to rhythmically entrain has been experimentally demonstrated in adults who have successfully synchronized with external rhythms over varying tempi (Snyder & Krumhansl, 2001) and who have synchronized to rhythms in pulses which were multiples or fractions of the original beat (Drake, Jones, & Baruch, 2000). Synchronization can therefore be seen as a naturally occurring cooperative human activity (Bispham, 2004; Brown & Volgsten, 2006).

Entrainment is a biological and physiological phenomenon which stems from sociality and culture. It occurs when two independent rhythmic patterns synchronize to a common periodicity (Bispham, 2004; Clayton, Will, & Sager, 2004). Entrainment is a naturally occurring human behavior which allows us to synchronize with an external rhythmic pattern and to time our movements with another during coordinated activities (Bispham, 2004; Clayton, Will, & Sager, 2004). Furthermore, it is a basic element of music. Rhythmic entrainment allows for the coordination of rhythmic behaviors with precise temporal synchrony (Merker, Madison, Eckerdal, 2009). Humans' ability to rhythmically entrain has been demonstrated in adults who successfully synchronized with external rhythms over varying tempi (Snyder & Krumhansl, 2001) and who synchronized to rhythms in pulses which were multiples or fractions of the original beat (Drake, Jones, & Baruch, 2000). Rhythmic entrainment operates off two assumptions. First, that internal oscillations, called attending rhythms, help create rhythmic expectations of future external events, and second, that the rhythms of the external events are what gives rise to the attending rhythms (Jones & Larger, 1999).

From an evolutionary perspective, the origins of rhythmic entrainment involved a social aspect. For example, during the synchronizing of fireflies, frogs, and fish (Gibson & Russell, 2006; Kotiaho, Alatalo, Mappes, & Pari, 2004). However, more recent research suggests humans to be unique in their ability to rhythmically entrain. Humans are the only species where both males and females can spontaneously synchronize body movements to acoustic rhythms (Patel, 2006). This is supported by the fact that rhythmic entrainment is found in all human cultures and throughout different musical contexts, such as singing and dancing (Kirschner & Tomasello, 2009). Furthermore, human synchronization is special because of its ability to be flexible and creative. Using the body, humans can entrain to complex rhythms and can adapt these movements in order to synchronize with changes in tempo. Such ability is especially important during dancing or musical performances when more than one person is involved (Keller, 2008).

1.2 Rhythmic Coordination

Rhythmic entrainment is the principle which allows humans to coordinate movements with one another (Merker, Madison, Eckerdal, 2009). It therefore facilitates interpersonal coordination. Performing coordinated actions strengthens rapport and facilitates social bonding between the individuals involved. These implications hold true even at an unconscious level (Isabella, Belsky, & van Eye, 1989; Bernieri, 1988). The positive relationship between interpersonal coordination and social bonding is due to a weakening of boundaries between the self and others. This in turn leads to information flow and knowledge transmission between the participating members (Wilson, 2001). Information flow and knowledge transmission occurs because there is a strong connection between perceiving and imitating another's actions. This connection is so strong that it is highly automatized. For example, imitation of perceived actions often occurs in social situations. Imitation results from the priming of a mental representation of a behavior (Byrne & Russon, 1998). Here, sensory input facilitates the perception of

other's behaviors. It then activates a mental representation of these behaviors and therefore increases the likelihood of these behaviors being performed (Wilson, 2001). One example of interpersonal coordination which requires imitation is joint coordination. Joint coordination often involves two or more people organizing and correlating rhythmic movements together. This type of coordination depends on the ability to predict another's subsequent actions and adapt promptly (Konvalinka, Vuust, Roepstorff, & Firth, 2010). Joint coordination therefore requires individuals to perceive and imitate others' actions.

Another type of rhythmic coordination, which is concerned with the relationship between perception and actions, is sensorimotor synchronization (SMS). SMS is the ability to rhythmically synchronize a motor response with a temporally perceived stimulus. Even with little or no formal training many people listen to music and respond by taping their feet, humming, and clapping. Music greatly depends on SMS for motor learning and motor performance. Musicians must precisely synchronize their productions with the audible and visual cues of others during a group performance (Jungers, Palmer, Speer, 2002; Palmer, 1997; Repp, 2005a). For example, group performers may heavily rely on auditory information to synchronize their performance, but they also rely on visual information. When auditory feedback is absent ensemble performers will increase the amount of visual information available by exaggerating gestures in order to facilitate accurate communication between performers (Goebl & Palmer, 2009).

A simple example of SMS is isochronous synchronization, which necessitates two main mechanisms. First, it requires one to produce a sequence of motor responses at the same frequency as the stimulus, and second, it requires the production of motor responses to coincide in time with the stimulus. In order to achieve these goals timing information from the stimulus must be extracted. SMS timing information is extracted using audition, but can also rely on proprioception information, especially in the absence of auditory feedback (Aschersleben & Prinz, 1995). The timing information is then internalized so motor outputs can be activated in advance of the stimulus and movements can be synchronized on the beat (Molinari, Leggio, Thaut, 2007). This often leads to a synchronization error (also known as negative asynchrony or anticipation tendency) (Repp, 2005), where subjects taps 20-50ms before the corresponding automated click. This phenomenon was first noted by Dunlap (1910) and has since been found in a number of studies (e.g., Aschersleben & Prinz, 1995, 1997).

1.3 Sensorimotor Synchronization

SMS between two individuals has not been well studied. This is because duet performances make it difficult to control the interactive behaviors used in coordination. Therefore, in order to observe participant's synchronization strategies studies have largely focused on single subjects. Individuals are typically asked to synchronize with a computer, which is used to simulate the behavior of a human partner (Repp & Keller, 2008). Some experiments request subjects to synchronize with an auditory sequence, then intentionally cause the sequence to slightly deviate from a perfectly isochronous pattern (Repp & Keller, 2004). This allowed researchers to identify two sources of sensory information that cause participants' motor activity to adapt to the deviations. The first is phase correction, where a change in the asynchrony adjusts the relative phase of the successive taps, but does not change the period of the underlying timekeeper (Goebl & Palmer, 2009; Repp, 2001). The second is period correction, where an adjustment to the internal timekeeper takes place. However, period correction only occurs when the tempo of the auditory stimuli changes (Repp, 2001; 2005).

Although single-subject studies of SMS give researchers more control over the elicited behaviors of the participants they disregard real world situations. For example, SMS often occurs with two more individuals, such as in instrumental group performances. In order to compromise with the level of experimental control and the number of participants Keller, Knoblich, & Repp (2007) investigated how pianists performed duets while playing with a recording of either themselves or with another pianist. They found that participants were better able to synchronize with a recording of themselves compared to that of another pianist. This encouraged Kovalinka, Vuust, Roepstorff & Firth (2010) to investigate interpersonal coordination between two individuals. They were interested in the mechanisms individuals use to synchronize with a given beat alone as compared to synchronizing with another person or computer. This was investigated by manipulating the amount of partner feedback each partner re-

ceived. Participants were asked to synchronize to feedback from either themselves, their partner, both themselves and their partner, or from a computer. The dyads made a mutual attempt to synchronize with each other, suggesting that in joint coordination, there is no leader-follower strategy. Instead, there is a continuous mutual adaptation to each other on a short millisecond timescale. This shows that coordination can be the result of a highly adaptive process, called "hyper followers" where individuals attempted to align each tap with the previous tap of their partner. For example, each member sped up their subsequent tap if their partner had been faster on the previous tap; the opposite occurred if their partner had slowed down. The dyad was also found to synchronize equally well when receiving feedback from both partners as compared to receiving feedback from only the computer. However, the dyad was worst at synchronizing when receiving feedback form only one partner. This suggests that it is possible to achieve equally good synchronization with a partner that is unpredictable but responsive (bidirectional condition), compared to a partner that is predictable but nonresponsive (computer condition). In the unidirectional condition, where feedback came from only one partner, the dyad would adjust their timing to the partner who could be heard. It is therefore most difficult to synchronize with a partner that is both unpredictable and nonresponsive. This suggests that successful interpersonal coordination is based on how well a member can accurately predict their partner's future actions and how promptly they can adapt to these actions.

Entrainment and rhythmic coordination have mainly been studied using behavioral experimental-paradigms. Although this has led to interesting insight about how we synchronize with external rhythms, and especially how partners mutually adapt to other, the neural mechanisms underpinning such coordination are poorly understood. This is evidenced by the fact that brain imaging research has yet to provide converging evidence identifying the neural correlates responsible for this phenomenon. Understanding the neural correlates of rhythmic entrainment are important as it will help us to understand how coordinated interactions unfold in time, and in turn, will make the process of joint coordination a less challenging task.

1.4 Duet drumming

Most research investigating rhythmic synchronization and entrainment have focused on individual's behavior using a finger-tapping paradigm. This paradigm implements a simple task where subjects are to tap with their right index finger in synchrony with an auditory sequence consisting of either tones or clicks (Repp, 2005). Although finger-tapping tasks are useful for examining individuals, musical activities often occur in a group setting, such as in ballet or in ensemble performance. These social situations are dynamic and individuals must be aware of others in order to coordinate their actions towards a common goal. Furthermore, music is a social activity and by contrast, finger-tapping is not a common activity engaged in by pairs or groups of individuals. Therefore, in order to focus on a more realistic sensorimotor synchronization task the current study has chosen to focus on a duet drumming task. Since ancient times drums have played an important part in music performance and recent research has suggested that coordinated drumming actions between two or more individuals facilitates pro-social commitment among those involved (Kokal, Engel, Kirshner, & Keysers, submitted). Therefore, joint drumming creates a natural environment for cooperative rhythmic tasks (Kirschner & Tomasello, 2009).

Psychological studies that have focused on cooperative tapping tasks have shown that seeing one another does not help to facilitate synchronization (Himberg & Thompson, 2009). Furthermore, visual information may actually hinder synchronization as many people choose to divert their gaze or cover their eyes in an attempt to block visual cues. This suggests that only auditory information is helpful for individuals synchronizing with each other (Repp & Penel 2004). However, we still have yet to fully understand how rhythmic entrainment is manifested between individuals in a musical performance. The question regarding how rhythmic entrainment arises and develops during a musical performance might be better addressed through examining the neurobiological basis of entrainment. With the advent of modern neuroimaging methods, it is now possible to investigate the neuro-biological basis of interpersonal rhythmic coordination. To the best of my knowledge, such a study has not been done before.

2 The Neural correlates of joint rhythm and action

2.1 Evolution

From a biological viewpoint it is interesting to consider the evolution of entrainment in animals. This is because a mechanism frequently thought of as uniquely human is rhythmic entrainment. However, entrainment has recently been demonstrated in other animals. For example. The cockatoo, 'Snowball,' who can dance and therefore synchronize to music. Snowball was able to entrain to the original and manipulated beat of his favorite song, *Everybody (Backstreet's Back),* by The Backstreet Boys (Fitch, 2009; Patel, Iverson, Bregman, & Schulz, 2009). As previously mentioned, a social context appears to be a common feature in animals known to synchronize. This includes the synchronizing of chorusing of crickets (Greenfield, 1994, 2005) and the antiphonal ensemble singing of plain-tailed wrens (Man, Dingess, & Slater, 2006). Through the study of entrainment in humans we may better understand the evolutionary origin of entrainment in animals. Therefore, another advantage to studying entrainment from a neurobiological basis is that it may shed light on the evolution of entrainment and rhythmic behavior. In this regard, the entrainment phenomenon has started to receive some attention. Entrainment constitutes three necessary elements: precise timing, synchronization, and joint performance. These three aspects have been studied by researchers independently and are described below.

2.2 Timing and movement

The brain imaging literature relevant to rhythmic entrainment has mainly focused on the neural correlates of time and movement. It was once thought that timing was processed in a specialized region of the brain, which was responsible for representing and processing temporal information (Xu, Liu, Ashe, Bushara, 2006). Furthermore, the cerebellum has been widely recognized for its role in controlling motor timing (Ivry, 1996). This led to the conjecture that the cerebellar cortex was the specific region of the timekeeping system (Molinari, Leggio, Thaut, 2007). However, it is now known that time is processed not only by the cerebellar cortex, but by a distributed neural network. This network specifies the cerebellum, frontal cortex, and basal ganglia as the primary structures involved in the processing of time. Furthermore, these structures interact to extract sensory information about time from the environment, and then pass this information to other brain regions for further processing. For example, timing information can be passed to the motor cortex in order to execute motor commands at appropriate times (Ivry, 1996).

The ability to time events is essential to the precise execution of movements (Müller et al., 2000) and it has been shown that different brain regions are activated depending on the timing of such movements (Harrington et al., 2000). Precisely timed movements have been shown to function through three independent interrelated neural systems. The first, the medial premotor loop, consists of the SMA, left caudate putamen, and the left ventrolateral thalamus (Alexander, DeLong, & Strick, 1986). This loop is responsible for the timing of movements, especially movements which are internally generated. The second, the right STG and the IFG, form a network for the retrieval and rehearsal of auditory information, particularly in the absence of external stimulation. The third system consists of the dorsal dentate nucleus and the SMC. This circuit is responsible for the processing of sensorimotor aspects in the execution of movements during timekeeping tasks (Leiner, Leiner, & Dow, 1995). It is therefore suggested that the cerebellum plays a large role in coordinating external and internal stimulus events. With connections between the dorsal dentate nucleus and the motor system the cerebellum is able to integrate multisensory information and use it to coordinate movements (Bloedel, 1992). With connections between the ventral dentate nucleus and the dorsolateral prefrontal areas the cerebellum is able to be involved in higher level cognitive processing, which is necessary for understanding rhythmic complexity and period correction in SMS tasks (Repp & Keller, 2004).

2.3 Rhythmic complexity

Just as with different timing of movements, different neural mechanisms may be recruited for different levels of rhythmic complexity. Using functional magnetic resonance imaging (fMRI) Dhamala and

colleagues (2003) identified brain regions which correlated with the complexity of rhythmic finger-tapping. They found that the rate of tapping correlated with ipsilateral activity in the pre- and postcentral gyrus located in the midbrain. This activation was a result of the additional internal effort needed to increase the rate of tapping after the initiation and execution of a motor command.

It was further found that with increased rhythmic complexity there was an increased involvement of the cerebellum and thalamus, and a simultaneous decreased involvement of the basal ganglia (Dhamala et al., 2003). Activations in the cerebellum and the basal ganglia were directly related to coordination of rhythmic movements as both are involved in motor control and action sequencing (Brown & Marsden, 1990; Casini & Ivry, 1999; Leiner, Leiner, & Dow, 1995). However, the decreased activity of the basal ganglia may suggest lesser involvement in sequencing complex movements (Dhamala et al., 2003). The specific roles of the basal ganglia are not well understood, but some suggest that it assists with the automatization of movement sequences (Brown & Marsden, 1990). The observed increase in thalamic activity is suggested to be a result of the recurrent connections it mediates between the basal ganglia and cerebellum to the cerebral cortex. These findings suggest that the cerebellum is associated with an increase in rhythmic complexity and that it plays a role in the timing mechanism of higher coordinative movements.

2.4 Neural correlates of synchronization

The finger-tapping paradigm frequently used to study SMS consists of two tasks: synchronizing with an automated metronome and continuing to tap in the absence of automated clicks. These two tasks are known to recruit different brain structures (Repp, 2005). fMRI research comparing synchronization and continuation tasks found the continuation task recruits the putamen, ventrolateral thalamus, and SMA. These neural substrates were involved in explicit timing. Furthermore, the IFG and STG were involved in the processing of auditory sensory memory (Rao et al., 1997). Using a similar design to that of Rao and colleagues (1997) a study examining temporal complexity found the SMA and basal ganglia to be activated in response to continuation tapping (Lewis et al., 2004). In addition to fMRI research, an EEG study by Gerloff and colleagues (1998) found the contralateral sensorimotor cortex and supplementary motor area to also be activated during self-paced finger-tapping. These results suggest that tapping without an external reference is a more demanding task than synchronizing with a given cue (Repp, 2005).

Just as different brain regions exist for different tapping tasks it has been suggested that different brain regions exist for phase correction and period correction in SMS. An Electroencephalography (EEG) study by Praamstra, Turgeon, Hesse, Wing, & Perryer (2003) suggest that there are neural distinctions between phase correction and period correction. For example, activation of the basal ganglia and the cerebellar loops with areas of the prefrontal cortex are associated with cognitive functioning. Since period correction is suggested to be a higher level of cognitive function (Repp & Keller, 2004), this implies that the neural substrates of period correction involve the basal ganglia and cerebellar - prefrontal cortex loops. For example, during an fMRI study participants were given a 'Visually Guided Task' where they were asked to move pegs from one side of a pegboard to the other. Later, they were given an 'Insanity Task' where they were to complete the same task, but with an unsolvable pegboard. Results showed that during the Insanity Task subjects displayed a large bilateral activation in the dentate nucleus that was three to four times larger than that of the Visually Guided Task. Also, the ventral portions of the dentate nucleus active during the Insanity Task innervated prefrontal regions of the cortex. This suggests that these regions of the dentate nucleus are involved in cognitive processes (Kim, Ugurbil, & Strick, 1994).

Activation of the basal ganglia and cerebellar loops with motor areas of the cerebral cortex are strongly associated with parameters of movement. This suggests that phase correction is mediated by the basal ganglia and cerebellar loops which have connections to motor areas. In one Position Emission Tomography (PET) study subjects learned eight-finger pattern movements. The brain activity of these movements was then compared during learning and during a performance after movements had already been learned. Results showed that a performance of previously learned movements activated the primary motor areas, including the motor cortex, medial regions of the cerebellar cortex, and dorsal portions of the deep cerebellar nuclei (Jueptner et al., 1997a; Jueptner et al., 1997b). Therefore,

these brain areas are associated with the cerebro-cerebellar circuits which innervate motor areas of the cerebral cortex and neo-cerebellar regions (Middleton & Stick, 2000). These findings demonstrate that individual loops do exist for different behavioral functions.

Insight into the cortical components responsible for the action and perception of SMS also derive from the finger-tapping paradigm. In a finger-tapping task the onset of the downward movement of the index finger should be preceded by cortical motor activity, for example, before a motor command (Hallett, 1994; Desmurget et al., 1999). After the initiation of a motor command there is a tap onset, where proprioception occurs and tactile information is extracted (Müller et al., 2000). During a magnetoencephalograhy (MEG) study examining cortical activation in an SMS task Müller and colleagues (2000) found three areas of the sensorimotor cortex to be activated during three distinct events. During a motor command the contralateral PMC peaked approximately 77 msec before the tap onset. This was interpreted as the preparatory and executive motor processes necessary for the initiation of a motor command. During propriception the contralateral S1 and the S1 inferior peaked at tap onset and approximately 75 msec after tap onset, respectively. The S1 and S1 inferior were found to function as two different input sources of sensory feedback. S1 was the input source for proprioceptive feedback during the downward movement of the finger, while the S1 inferior was the tactile feedback from the pad of the finger after tap onset. These results suggest that evaluations of 'keeping time' occur through sensory feedback processes and are represented by a late component localized in the S1 inferior. This study further demonstrates the important of the sensorimotor cortex in the evaluation of timing.

2.5 Neural correlates of practice

Neuroimaging studies have also been interested in the effects of repeated practice on the brain activity. This is because after practice differences in neural activation between new and previously learned tasks decrease, suggesting that learned tasks require different neural mechanisms (Mayville et al., 2002). Lewis and colleagues (2004) therefore investigated the brain activity associated with an over-learned movement task. Results showed that during the initiation of movements and synchronization tasks activity was found in the frontal and prefrontal brain regions. However, when participants were to continue tapping with an external cue this this pattern of activation did not show. This suggests that the continuation tasks only required overlearned motor commands to be executed. This further suggests that once movements have been selected and initiated they do not require processing in the frontal and prefrontal cortices because the over-learned tasks have become automated and therefore require less attention. Therefore, once an over-learned movement sequence has been selected and initiated, the ability to continue this sequence does not depend on the temporal complexity of the rhythmic pattern (Lewis et al., 2004; Vorberg & Wing, 1996).

2.6 Neural correlates of joint action

A large body of research has suggested that an individual is able to understand the actions of others (Puce & Perrett, 2003), and that viewing these actions can activate motor and somatosensory neurons as if they were one's own actions (Keysers et al., 2004). This concept stems from research concerning mirror neurons in the brains of monkeys which were found to respond equally to the execution of a performed action within the monkey itself and during the observation of the same action performed by an experimenter (Fogassi et al., 2005)Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992).

Recent human studies have supported the existence of a mirror neuron in the ventral and dorsal premotor areas, the IPL, the MTG, the SMA, and the somatosensory cortex, which is collectively referred to as the mirror neuron system (MNS). Gallese (1998) and Rizzolatti and colleagues (2001) suggest that mirror neurons help us to directly map our observations of others' actions onto our own motor system. The MNS has also been proposed as a fundamental element which allows us to engage in joint actions (Newman-Norlund et al., 2008; Sebanz, Rebbechi, Knoblich, Prinz, & Frith, 2007).

Joint performance involves listening and coordinating as well as precise timing for the execution of motor commands. Indeed, the literature suggests that joint performance activates several brain regions including the temporal, auditory, and visual cortex. Recent research by Newman-Norlund and

colleagues (2008) found that activation in the IFG was larger when participants balanced a ball with another as compared to when they balanced the ball alone. This demonstrates the role of the MNS in joint actions. Another experiment with similar findings demonstrated greater activity in the right IFG and the bilateral ipsilateral parietal-lobule during the planning of complementary actions as compared to imitative actions (Newman-Norlund et al., 2007b). However, the IFG and the parietal lobule may not be part of the MNS as they include neurons which are not mirror neurons (Kokal, Gazzola, & Keysers, 2009; Thioux et al., 2008).

In order to identify the circuitry directly involved in joint action and to examine the role of the putative mirror neuron system in joint action Kokal, Gazzola, & Keysers (2009) conducted an fMRI experiment where participants were asked to cooperate with an experimenter to shape two sticks on a response box in either a straight line or an angle. Results showed a distributed neural network to increase in activity during the joint action task as compared to the solo observation and solo action tasks. This included the posterior parietal lobe (inferior and superior parietal lobule; precuneus), the basal ganglia, and the cerebellum. These brain regions are known to be involved in learning and using novel sensory-motor associations (Gold & Shadlen, 2007; Hikosaka et al., 2002). Furthermore, this study suggests that joint action contains a dual process. One set of networks, includes the MNS, transform observed actions into representations of similar actions. This network functions to help integrate our own actions and the actions of others during joint actions tasks. For example, this network may help the viewer to understand and predict the goals of observed actions from another (Umilta et al., 2001). This allows the observer to choose an appropriate action which has been primed. The other network includes regions of the prefrontal cortex, the posterior parietal lobe, and the temporal lobe, which are adjacent to the MNS. This network's role is to choose the most appropriate action and then to integrate our own actions with the actions of others.

2.7 Current study

Most research investigating the neural correlates of rhythm focuses on the individual and are therefore not specific to joint action. However, it has been suggested that more brain regions are activated during a joint performance as compared to a solo performance (Newman-Norlund et al., 2008). Furthermore, the few studies which do take into account joint action are not specific to rhythmic coordination. And no study, to the best of my knowledge, examines rhythmic coordination during a joint-action rhythm-task. Therefore, in order to account for group music making this study focuses on a joint drumming task where pairs of participants are asked to mutually respond to each other. Furthermore, in order to simulate a more realistic situation with joint action we use a drumming task where partners are to drum using wooden claves. Claves are a genuine instrument used in the percussion section of bands and orchestras. Using an actual musical instrument therefore encourages participants to play musically during the experiment. This facilitates engagement from the participants and helps yield results that are more reflective of a true collective time keeping task. Also, a majority of drumming experiments have been behavioral based studies, which leads to a gap in the literature concerning the neural correlates of drumming. And although there are brain imaging studies which focus on individuals' rhythmic productions, the neural correlates of joint coordination has not been fully investigated. Lastly, some research disregards the musical ability of trained musicians by considering the term 'musician' to be a socially constructed definition. However, studying trained musicians can be seen as an advantageous as they are able to learn rhythms more quickly and are more comfortable with complex rhythms. The current study is therefore interested in how partners mutually adapt to each other's timing. This task is easier to isolate if the participants can learn the rhythms well and understand the instructions quickly. Also musicians are used to adapting to one another during performances, which may help to ensure that they are aware of their partner, even if they adapt to each other on an unconscious level.

The current study is therefore interested in how partners mutually adapt to each other's timing and the neural correlates involved in such actions. Understanding mutual adaptation will be easier if the participants can learn the rhythms well and understand the instructions quickly. For this reason, the current study aims to used musicians. Also, musicians are familiar with adapting to one another during performances, which may help to ensure that they are aware of their partner, even if they adapt to each other on an unconscious level. The current study therefore aims to compare musicians

with a human partner, alone, or with a human-generated prerecorded auditory sequence. During these drumming conditions simple (synchronized) and complex (asynchronized) rhythmic patterns will be investigated. Furthermore, three different rhythms will be employed, which are isochronous, double, and long-short-short (LSS). The questions guiding this research are as follows: What are the neural substrates involved in a joint drumming task as compared to drumming alone or with a computer? Are patterns of neural activation different for simple and complex rhythms? And finally, how well are partners able to adapt to each other's timing during a joint drumming task?

In this study, pairs of participants first took part in a behavioral task, which was used to simulate that of the fMRI study to be conducted at a later time, and in order to familiarize them with the rhythmic patterns and procedure. Participants then took part in the brain imaging study with the same partner they were with in the behavioral task. First, one partner was brain scanned, while the other partner performed in a separate room. At the end of the procedure the partners switched places. Participants were instructed to perform the same rhythms in the same order as in the behavioral study. During the brain imaging study participants were audio recorded so that the levels of synchrony and asynchrony could be analyzed later.

The overall aim of this behavioral and fMRI study is to examine rhythmic synchronization, and the neural correlates of such synchronization, between human partners, solo performances, and joint performances consisting of a human and a human-generated prerecorded audio track. This behavioral study is particularly interested in the interaction effects of the various conditions and their effect on the variability of synchronization levels among participants. Because there are less timing demands placed on participants during the Simple condition, where performers are to match there onset times with their partner's timing, synchronization variability is expected to be the lowest for the Simple condition as compared to the Complex condition. Furthermore, performing with a human partner is expected to facilitate the creation of a shared representation of performance (Kirschner & Tomasello, 2009). Therefore, the lowest synchronization variability is expected to result from the Human-Human condition, as compared to the Solo condition and the Human-Computer condition. Also, there are the least timing demands placed on participants during the isochronous rhythm because there is only one rhythmic duration (eighth note) involved in this condition. The LSS rhythm contains the most timing demands as there are two different rhythmic durations (an eighth note and sixteenth note) and three onset times in this condition. And the double rhythm contains timing demands which are more than the isochronous rhythm, but fewer than the LSS rhythm. This is because it contains two different rhythmic durations, but only two onset times. Therefore, the isochronous rhythm is expected to have the lowest synchronization variability, LSS is expected to have the highest synchronization variability, and the double rhythm is expected to be in between these other two rhythms. These hypotheses result in the expectation that the Simple Human-Human Isochronous condition will contain the lowest synchronization variability compared to all other conditions.

The particular aim of the fMRI study is to identify the neural substrates activated during joint rhythmic coordination. Precisely timed movements are mediated by a number of brain regions, including the cerebellum and the SMC, which are thought to process sensory information in order to execute movements (Bloedel, 1992: Leiner, Leiner, & Dow, 1995). Increased activation of the cerebellum has also been shown during increased time demands during more complex rhythms. Furthermore, the SMC and PMC, areas of the motor cortex, are known for planning and executing motor movements (Dhamala et al., 2003). Therefore, the cerebellum, SMC, and PMC are expected to increase in activation during Complex condition tasks as compared to Simple conditions. These brain areas are expected to show the most activation during the conditions with more rhythmic complexity as these conditions will require more planning and executing of movements. Furthermore, during joint actions, as compared to solo action tasks, the posterior parietal lobe, the basal ganglia, and the cerebellum were shown to increase (Kokal, Gazzola, & Keysers, 2009). These brain regions are involved in learning and sensorimotor activities (Gold & Shadlen, 2007; Hikosaka et al., 2002). Therefore, the posterior parietal lobe, basal ganglia, and cerebellum are expected to increase in activation during the Human-Human conditions as compared to the solo and human-computer conditions. Lastly, SMS studies have shown that continuation tasks recruit the putamen, ventrolateral thalamus, SMA, and basal ganglia. These areas are known to be involved in timing demands and during tapping tasks without an external cue of

the beat (Lewis et al., 2004; Repp, 2005). It is therefore hypothesized that the putamen, ventrolateral thalamus, SMA, and basal ganglia will show increased activation in the Solo condition compared to the Human-Computer condition.

3 Techniques

3.1 Use of fMRI

fMRI is a brain mapping technique developed in the 1980s by modifying the magnetic resonance imaging (MRI) technique. fMRI detects differences in the magnetic properties of hemoglobin when it changes from oxygenated to deoxygenated blood (Kwong et al., 1992). During resting conditions cerebral blood flow (CBF) and cerebral blood volume (CBV) are regulated through neuronal activity. However, due to metabolic responses the rising of CBF and CBV uncouple from oxygen consumption. This uncoupling results in a decrease in deoxygemoglobin and therefore provides the blood-oxygen-level dependent (BOLD) signal used indirectly by fMRI studies (Belliveau et al., 1992). Figure 1 shows the time course of CBF and CBV within the hemodynamic response.

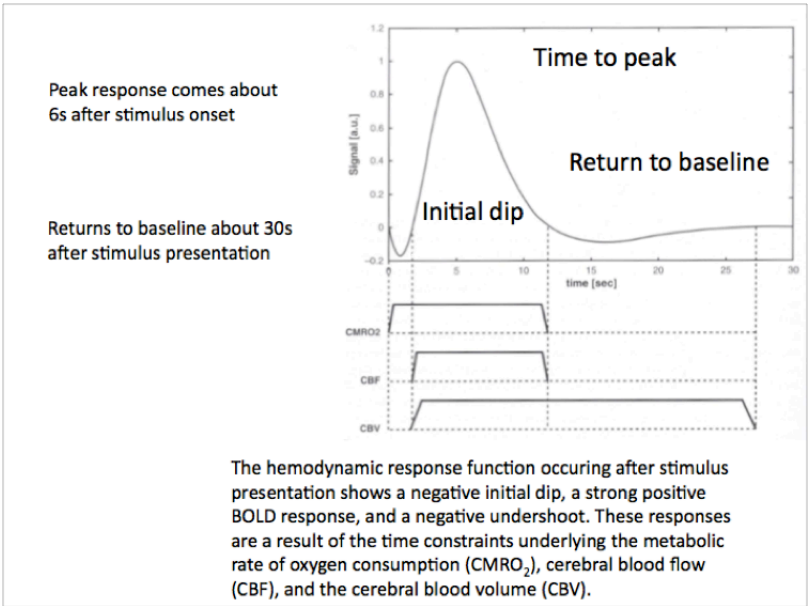


Figure 1: the hemodynamic response function (HRF)

fMRI has advantages over other brain imaging techniques. It is a noninvasive technique, which records brain signals without the risk of radiation exposure seen in other methods, such as in X-ray Computer Tomography (CT) and PET scans. It also has a high spatial resolution of approximately 2-3 mm. Lastly, it records brain signals from the entire brain instead of only from cortical surfaces, such as in EEG and MEG.

The choice of using fMRI was based on the resources available at the University of Sheffield as the university owns the machinery necessary to conduct fMRI experiments. Also, several professors within the psychology and radiology department have expertise in the neuroscience field. For example, Dr. Ian Wilkerson and Dr. Lawrence Parsons, who have collaborated on several experiments, demonstrating their ability to work together. Furthermore, Dr. Parsons has been conducting research on the cognitive neuroscience of music for over a decade, suggested that his expertise is invaluable in the field. With a background in music I am particularly interested in the neuroscience of music and therefore felt this brain imaging study to be an ideal experiment for me to take part in.

3.2 Physiology underlying the BOLD signal

When a task is performed the neurons involved begin to fire. This firing then leads to a decrease in ionic gradients within the neurons because of the release of neurotransmitters at the synapse. Therefore a source of energy is required to re-establish the gradient in the neuron. This energy is supplied by the conversion of adenosine diphosphate (ADP) to adenosine triphosphate (ATP). The source of energy required for this conversion is the metabolism of oxygen and glucose within mitochondria of brain cells. The blood stream is the transport system used to supply glucose and oxygen to the mitochondria for the ADP-ATP conversion. The conversion occurs in two steps: (1) anaerobic glucose metabolism (glycosis), and more importantly, (2) the transcarboxylic acid cycle (TCA). This cycle particularly requires large amounts of oxygen. Two by products of the TCA cycle are carbon dioxide (CO_2) and water (H_2O), which leave the TCA cycle by diffusing back into the blood stream.

During neuronal activity ATP usage increases meaning the TCA cycle produces more CO_2 and H_2O. When this happens blood becomes more deoxygenated, which is compensated for by the dilation of blood vessels. Dilation occurs either because CO_2 is a vasodilator, or because neurotransmitters release nitrous oxide (N_2O), which is also a vasodilator. The exact mechanisms underlying vasodilation is an area currently under research. However, it is agreed that neurovascular coupling results (Belliveau et al., 1992). This increase in blood flow causes an increase in blood volume and therefore causes blood oxygenation to increase to point where oxygenated blood outweights deoxygenated blood two to one. Figure 2 diagrams the changes in blood oxygenation levels due to neuronal activity.

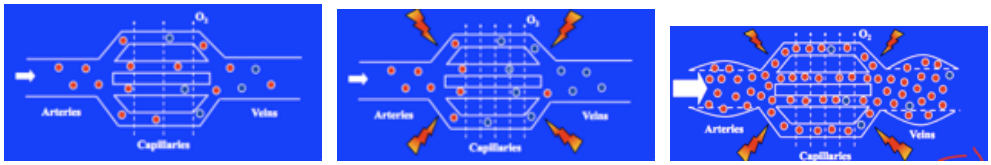


Figure 2: shows the blood vessel at baseline. As neuronal activity increases oxygenated blood is used, leaving only deoxygenated blood behind. This causes vasodilation and subsequently an increase in blood flow.

3.3 Sparse sampling technique

The current study involves listening and responding to auditory stimuli, which is complicated by scanner noise. For this reason the current study implements a more recent method for acquiring brain imaging data called The Sparse Sampling technique. The sound created from the MR scanner is known to interfere with the perception of auditory stimuli because the flow of electric currents creates a magnetic field, which surrounds the conductor. The conductor is then in a magnetic field, causing the Lorentz force to act upon it. This is what happens to the gradient coil of an MR scanner which produces the transient magnetic fields used for slice selection and readout. This means that the Lorentz force deforms the gradient coil, and subsequently spreads to the surrounding air. When this deformation spreads across air molecules it generates a noise. This is noise is analogous to the sound produced when a membrane of a speaker is displaced (Mansfield et al., 1998). The amount of noise given off by the gradient coil is proportional to the field strength of the scanner (the greater the field strength, the louder the noise) and is typically over 100 dB SPL (sound pressure level) (Price et al., 2001). This means participants must use earplugs (and sometimes headphones) in order to avoid damaging their hearing. Although measures are taken to minimize scanner noise, the excess sound from the scanner's echo planar imaging sequences still interferes with the perception of auditory stimuli and causes significant changes in response to these stimuli during experimentation (Herrmann et al., 2000; Novitski et al., 2001, 2003).

Not only does scanner noise interfere with auditory perception, it also results in a Blood Oxygenation Level Dependent (BOLD) response. Auditory discrimination tasks with MR scanner noise have shown the BOLD response of the auditory cortex to be modulated (Shah et al., 1999). Comparing images taken after a period of scanning with those taken after a period of silence has revealed the PAC to be activated in response to scanner noise (Bandettini et al.,1998). This noise has also been a

problem for visual experiments, showing the BOLD response of the visual cortex to decrease by 50% during scanner noise (Cho et al., 1998). Investigation into the time course of the scanner noise, and the resulting activation of the primary and secondary auditory cortex, has determined that these brain regions peak about 4-5sec after the onset of the stimulus, and then decay about 5-8secs after they peak (Hall et al., 2000). This time course suggests that experimental designs with auditory stimuli can reduce the effects of scanner noise by implementing a longer repetition time (TR) of approximately 9-13sec (Hermann, Brechmann, & Scheich, 2010).

A solution to the problem of scanner noise and its effect on brain activity is to present stimuli during the silent periods of the scanning procedure. This is when the scanner is on, but not scanning a brain volume. This is achieved by increasing the interscan interval. This means that auditory stimuli is interleaved with the MR scanner noise and is presented during periods of silence (Hall et al., 2000). Because the BOLD response lags behind the electrical brain response by approximately six seconds there is time to present stimuli during the period of silence. The lagging BOLD response is then recorded approximately six seconds after the presentation of stimuli. Although the recordation of the BOLD response, known as the acquisition of brain volumes (TA), means that the MR scanner will produce noise, it is seen as less disruptive compared to scanner noise present throughout stimuli presentation. This procedure is known as sparse sampling (Hall et al., 1999) or clustered acquisition (Edmister et al., 1999). It is a relatively new brain imaging technique, which has been used in a number of auditory experiments and has proved effective towards relieving participants of excess auditory modulation and stress, while still successfully detecting changes in the hemodynamic response function (Hall et al., 1999; Hunter et al., 2006; Okada & Nakai, 2003; Zaehle, Wüstenberg, Meyer, Jäncke, 2004).

Our approach to collecting data using the sparse sampling method was through using a block design task where two conditions (task, rest) were alternated. Using this design the hemodynamic response function (HRF) was comprised of three stages: 1) an initial stage where signal increases, 2) a plateau stage, where signal temporarily peaks, and 3) a final stage, where signal decreases back to its baseline level. Because the initial and final stages of the HRF take more than five seconds to increase or decrease, a second-by-second measurement of this change is unnecessary. The maximum difference in this signal change can be measured by acquiring single volumes of brain images during baseline conditions (rest) and at the end of the stimulus presentation (task). (Okada & Nakai, 2003). Figure 3 compares the time course of volume acquisition between the continuous sampling method and the sparse temporal sampling method. It demonstrates how the sparse sampling method does not result in a loss of information. Figure 4 demonstrates the procedure of the sparse temporal sampling method as it relates to the HRF, showing that information is acquired during peak and baseline levels of the HRF.

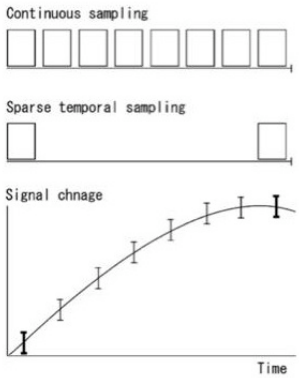


Figure 3: Comparison of two fMRI sampling methods in a block design task. Continuous sampling acquires data at all time course changes (shown by thin bars). Sparse temporal sampling acquires data only during baseline and activation plateau phase (shown by thick bars). The boxes mark the acquired volume in each TR period.

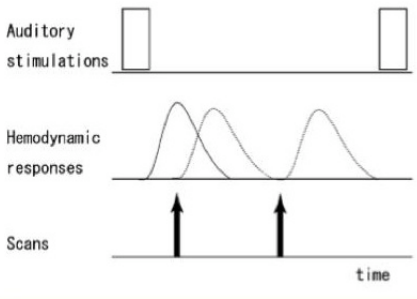


Figure 4: Depiction of time course used in the sparse temporal sampling method.

When comparing sparse sampling and continuous sampling techniques it is advantageous for experiments with auditory stimuli to use sparse sampling for several reasons. First, using sparse sampling means that the effects of scanner noise on brain activation, particularly in the auditory cortex, are minimized and therefore help to ensure that the fMRI data obtained is a result of the auditory stimulus, not of the scanner noise. Secondly, with sparse sampling the MR signal-to-noise ratio is increased. Although one may argue that in sparse imaging, compared to continuous imaging, there are less data samples, which results in a decrease in statistical power, it is argued that this is compensated for with the increase in the signal-to-noise ratio. Hall et al (1999) tried to compensate for the decrease in statistical power by increasing the number of acquired volumes. This method was termed a 'quasi-sparse imaging method.' In such a method scanner noise was kept to a minimum by acquiring several volumes at the end of each epoch. This method was compared to sparse imaging to see which method produced the most significant results. The quasi-sparse method was significantly worse at detecting auditory activation. The third advantage of using a sparse sampling method is that the loud noise originating from the scanner is diminished during data acquisition. During a task involving auditory stimuli scanner noise interferes with the stimuli and therefore changes the participants' perception of them. Participants also find the noise distracting and stressful, which impairs the ability to concentrate and subsequently decreases task performance (Hall et al., 1999; Smith, 1988). Although the auditory system is able to localize sounds, determine sound sources, and separate auditory information from noise, this is difficult to do with noises that are extremely loud because the sounds blend together. In fact, stimuli have been reported to become indistinguishable from noise, a phenomenon known as auditory perceptual grouping (O'Conner & Sutter, 2000; Okada & Nakai, 2003). Using a sparse sampling method therefore increases the external validity of auditory experiments, as participants are better able to hear to the stimuli and better able to concentrate.

3.4 Analysis technique: GLM

The general linear model (GLM) was used to analyze the data. GLM works by fitting a model to the fMRI data and it assumes that at each voxel the fMRI changes fit linearly to experimental effects (Friston et al. 1994). A GLM consists of a set of equations which express a predicted time course of the fMRI signal as a weighted sum of linear terms. These terms then represent the effects of interest and confounds. Therefore, GLM is an expression which assumes that fMRI changes are linearly related to the experimental effects at each voxel. It is common in fMRI studies to lack hypotheses about the relationship between experimental effects and fMRI changes. Under such circumstance GLM is estimated for each voxel and the results are known as estimated weights. The variability of these weights across voxels then creates the fMRI localization information. The fMRI signal at any voxel is expressed as a vector of successive observations, termed y_i, $i = 1 \ldots N$, where N equals the total number of observations. In fMRI experiments this equals the total number of scans. Each observation, y_i, is then modeled as a weighted sum of X_k ($k = 1 \ldots K$) "regressors," where K is the number of effects of interest represented in the model, and the residual (zero mean) error:

$$y_i = \beta_1 x_{il} + \beta_2 x_{i2} + \ldots \beta_k x_{ik} + e_i \quad (1)$$

where β_j are the unknown weights for each K effect. ß, the unknown weights, are the model parameters. In matrix form the expression now becomes:

$$Y = X\beta + e \tag{2}$$

Figure 5 outlines the mathematical frame work of the GLM. In the matrix X represents the modeled effects and is termed the design matrix. The aim of the GLM then is to estimate the model parameters, vector β, at each voxel.

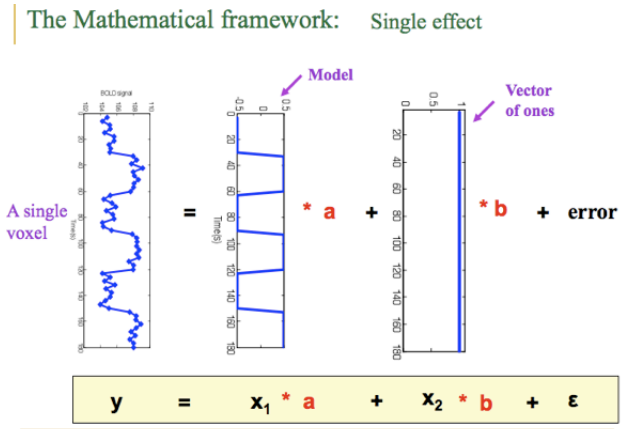


Figure 5: Depiction of the mathematical framework used in the GLM (Courtesy of Ying Zheng).

Regressors are either effects of interest, such as conditions, or effects of no interest, such as confounds or constants. Regressors of one category must be distinguished from the other in order to separate signal from noise. For example, body motion is a regressor of no interest and contributes to a decrease in the signal to noise ratio. However, regressors of interest model the part of the fMRI signal which relate to the experimental conditions (i.e. simple versus complex rhythms). This type of regressor is obtained by convolving boxcar functions, mathematical representations of events of interest, with a model of the event-related fMRI response, such as the HRF or FIR (finite impulse response). Choosing a representation for the conditions (tasks, stimuli) is dependent upon the assumptions given to the relationship between the stimuli (when present), the neuronal response, and the hemodynamic signal. Here, we assumed a box car shape with a time window equivalent to the canonical HRF. The regressors of no interest for example, structured noise or drifts in scanner sensitivity, are presented in order to model the remainder of the fMRI signal. The inclusion of these effects are important because they allow for a more reliable estimation of the GLM coefficients and therefore increase the confidence level of the results (Lund et al. 2005).

The next step of GLM is to estimate the unknown parameter, vector β. This is done through the method of ordinary least squares, which aims to minimize the residual sum of squares. That is, the sum of the squared difference between the predicted signal change and the actual signal change. This guarantees maximum likelihood estimates, meaning that the estimated parameter values are the most likely to occur given the data. Next, within the GLM, the amplitude of β (or a combination of coefficients) is tested to see if it significantly differs from zero. For example, where β_1 is a task and β_2 is rest, a contrast $(\beta_1\text{- }\beta_2)^2$ is first created then a t-test is applied to it in order to assess its significance:

$$T = (\beta_1 - \beta_2)/\text{std}(\beta_1 - \beta_2)\ (\beta_1 - \beta_2)/S(\beta_1 - \beta_2) \tag{3}$$

Where std(ß) is the estimated standard deviation of β, which is proportional to the sum of squares of the residuals, S(β). Knowing this, the following t test is used in order to test the hypothesis, $\beta > 0$:

$$\beta/\text{std}(\beta) \tag{4}$$

Such a contrast is commonly used to test if a predicted experimental effect (ON-OFF stimulus) is likely to be in the data.

Another test commonly used is the F test. This tests whether there is at least one coefficient within a set of many coefficients that is different from zero. Such a test is used when the task of interest is modelled by several regressors. For example, when testing for deviations in the hemodynamic response compared to the canonical HRF, it is useful to include the derivative of the HRF in the model, along with the canonical HRF, in order to map brain regions where the BOLD signal can be represented by a linear combination of the canonical HRF and its temporal derivative. This is accomplished using an F test. An F statistic is created by examining the full model and the reduced version of the model that is obtained if the hypothesis is true ($\beta_1 = 0$). Showing the reduced model by the subset of parameters, β_2, the F statistic is:

$$\frac{\left(\frac{\mathrm{S}(\beta_2)-\mathrm{S}(\beta)}{p-p_2}\right)}{\left(\frac{\mathrm{S}(\beta)}{\mathrm{N}-p}\right)} \tag{5}$$

where, S(β) is the squares of the residuals for the full model and S(β_2) for the reduced model. Furthermore, p is the rank of the design matrix for the full model, and p_2 for the reduced model. Lastly, N is the number of scans.

These statistics result in 3D maps showing the t or F statistic of the effect of interest at each voxel. The statistics are derived for each voxel independently; however, the significance of the deviations from the null hypothesis (conversion of t statistics to Z scores) must be assessed in relation to the maximum score or the size of an activated brain region. Only the false-positive rate (type I error) can be controlled; type II error control is typically ignored as it is too difficult to specify a quantitative alternative hypothesis in fMRI experimentation. In other words, the lack of activation does not imply that there is no underlying brain activity (Bénar, Bagshaw & Lemieux, 2010).

Several methods exist to calculate significant thresholds for fMRI experiments. The most important method is based on the theory of random Gaussian fields (Worsley et al. 1992). This parametric framework controls for the fact that with a null hypothesis with no significant effect, the maximum value of the map can be above the threshold with a p value of α, where $\alpha = 0.05$. This threshold is corrected for the large multiple comparison problem which arises when testing a large number of voxels at the same time, and corresponds to the family-wise error rate (FWE). A second method for calculating significance thresholds is to use a threshold based on the extent of a given cluster of activated voxels above an uncorrected threshold with a low p value (usually $\alpha = 0.001$) (Poline et al. 1997; Cao, 1999). The thresholds obtained with random field theory are usually high and require a large level of image smoothing for example, with a filter FWHM of 10-15mm (Bénar, Bagshaw & Lemieux, 2010).

4 Method

4.1 Participants

In this study, we recruited eight healthy volunteers with no reported history of neurological or psychological disease and had no hearing impairments. These volunteers were musicians, defined as music degree-seeking students of the Department of Music at The University of Sheffield. They actively played an instrument or sang, including ;cello, oboe, violin, and piano. Six of them were undergraduates while two participants were PhD students. All participants were in their mid-twenties. All volunteers gave informed consent and had the right to withdraw at anytime. We did not attempt to recruit participants based on racial or financial background. We obtained ethical approval from the Department of Psychology of the University of Sheffield.

4.2 Drumming Stimuli

A professional drummer, Gareth Widdowson, a PhD student at the University of Sheffield, prerecorded the stimuli. The recordings were intentionally created to have slight deviations in onset timings and

intensities in order to replicate the drumming of a human. This was done so that the recordings would not sound mechanical and therefore required participants to adapt and respond to the deviations.

4.3 Materials

Two sets of wooden claves for drumming were purchased from Percussion Plus (Reading, UK). One set was in the treble clef range, the other was in the bass range. Other materials include two sets of headphones so the participants could hear their partner's performance.

4.4 Experimental tasks

The experiments took place over a period of one week in the Academic Unit of Radiology at Royal Hallamshire Hospital at The University of Sheffield. Two experiments were conducted on 27 January, from 13.00-16.00, and from 16.00-19.00. The other two experiments were conducted on 3 February, at times corresponding to the first two experiments. These experiments took place.

All participants performed a joint drumming task and a solo drumming task. The joint drumming task consisted of blocks one through three, (i.e. Isochronous, Double, Long Short Short), and each block consists of six trials. The first three of these six trials, termed "SIMPLE" required each pair to drum together in synchrony. The last three of the six trials are termed "COMPLEX" and required each pair to drum in an asynchronous interlocking fashion. In the complex condition each partner performed the same rhythmic pattern, only that the rhythmic pattern for one of the partners was shifted by either one eighth note or by one quarter note so that the partners pattern's were out of phase.. Furthermore, trials one two, four, and five had a human-human ensemble arrangement where both participants played together (either in synchrony or asynchrony). Trials three and six had a human-computer ensemble arrangement, called play-along, where the scanned participant played only with the prerecorded stimuli. During play-along trials the unscanned participant sat and remained quiet.

The solo drumming task was the fourth and final block. Here the scanned participant was given a 15s computer introduction. After this introduction the stimuli ceased while the participant was asked to continue drumming alone. Participants were instructed to start drumming the scheduled rhythmic pattern as soon as possible and to continue drumming even after the computer stopped. The participant went through six trials during the solo block, which mirrored the rhythms of the previous blocks (e.g. isochronous simple, isochronous complex, double simple, double complex, LSS simple, LSS complex). Fig. 6 is the instruction sheet given to each participant during the experiment. It shows each rhythm in musical notation and provides a chronological diagram of the drumming tasks.

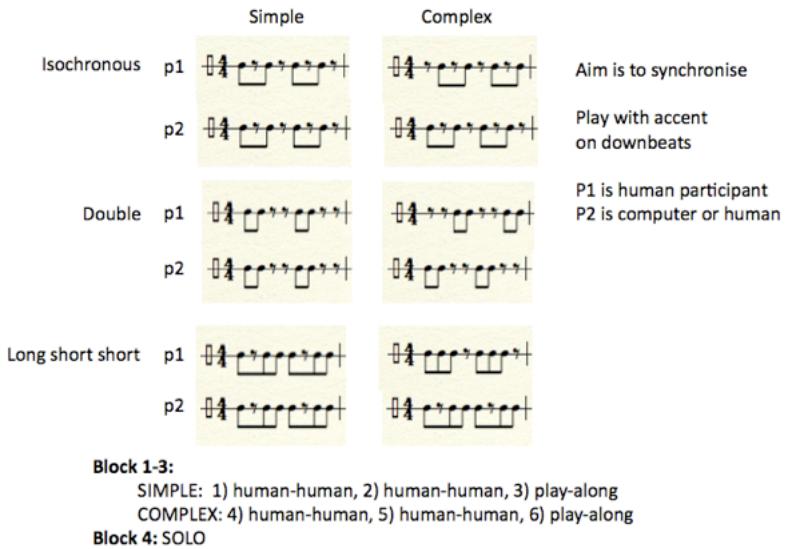


Figure 6: Joint Drumming Instruction Sheet

4.5 Procedure

For the drumming task participants were first organized into pairs. This was done so that each participant had only one partner. In this way, one participant would be scanned while their partner was sitting in a separate room, which was located just behind the scanner. After the task and scanning was completed for one participant the partners would switch positions.

Before the fMRI experiment participants were familiarized with the drumming tasks through a behavioral experiment, which simulated the fMRI experiment and served to trained the participants with the correct rhythms and synchronization conditions. The behavioral experiment consisted of the same procedures and materials, without the scanner, and took place in the Department of Music at the University of Sheffield.

4.6 fMRI scanning

Whole head functional MRI scans were performed using echo-planar imaging (EPI) on a 1.5T Eclipse system (Philips Medical Systems, Cleveland, OH) at the Academic Unit of Radiology at The University of Sheffield using a standard quadrature radiofrequency transmit/receive head coil (Hunter et al., 2006; Wilkinson et al., 2005). A list of fMRI parameters used in this study is described in Table 1.

Parameter	**Value**	**Notes**
Slice Thickness	4mm	Thicker slices give better SNR, but lower spatial resolution
Functional Slices	35	Number of slices in one brain volume
Field of View	240mm	Head has to fit in the space as FOV is the size of the 3D spatial encoding area of the image.
In-plane Matrix	240 X 140	Usually X^2
Voxel Size	4mm	Voxel size = FOV/matrix; bigger voxels give better SNR but poorer spatial resolution
Time to Repeat (TR)	15s	The number of slices per volume that can collect is determined by TR and matrix size
Volumes (Times Points)	30	Limited by scan duration
Time to Echo (TE)	35ms	Time between applying the 90° pulse and the peak of the echo signal in spin echo and inversion recovery pulse sequences
Acquisition Time (TA)	4s	Time required to collect the image data
Scan Duration	488s	Scan duration = volumes * volume acquisition time
Session Duration	1 hour	Limited by subject motivation and expectations. (More is not better if subjects get bored/sleepy).
Flip Angle	30˚	Amount of rotation of the magnetization vector produced by an RF pulse, respective to the direction of the magnetic field. Formed by apply a gradient, then reversing the direction of the gradient.

Table 1: fMRI Parameters Used in Current Study

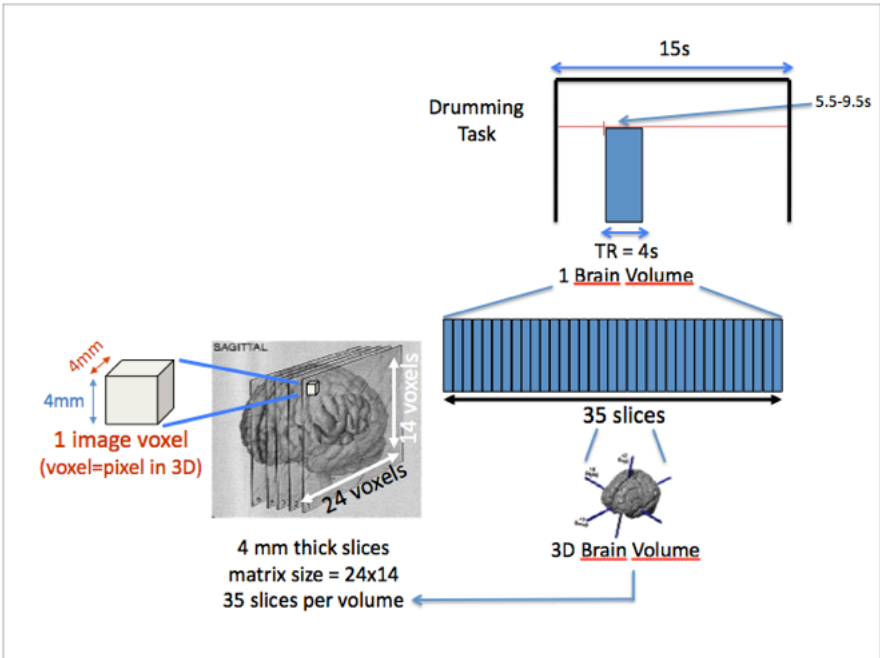


Figure 7: Illustration of the parameters used in the current study.

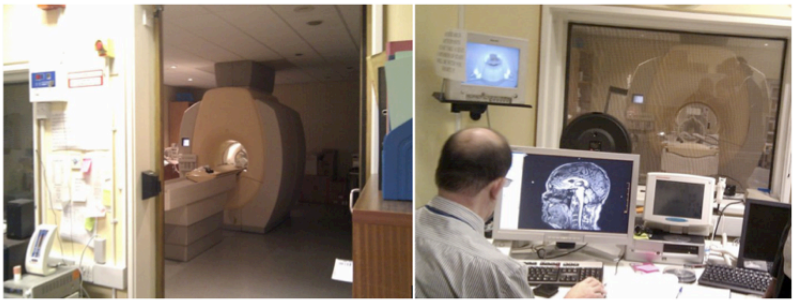


Figure 8: Pictures of fMRI scanner and laboratory used to conduct this study, which is located in the Radiology Unit of The Royal Hallamshire Hospital at The University of Sheffield (Courtesy of Maria Panagiotidi).

4.7 Data Analysis

Before data analysis it was necessary to discard one or two dummy scans from the beginning of each time series. This is useful because of the large signal change which occurs within the first one or two images of a functional dataset as tissue reaches a steady state of readiofrequency excitation. It is typical to discard approximately 15 s worth of data. Here we discarded between 4-8 s of data. Functional images were created by generating statistical parametiric maps (SPMs). They were then preprocessed using SPM8 software (Statistical Parametric Mapping, Wellcome Department of Imaging Neuroscience, University College London, London, http://www.fil.ion.ucl.ac.uk/spm/) implemented in MATLAB 7.10 (Mathworks, Inc., Sherborn, MA, USA). SPM is a software toolkit designed to analyze functional imaging data, including fMRI data. Each subject's image was realigned to account for head motion, then normalized using the template of the Montreal Neurological Institute. The scans were also smoothed using a Gaussian kernel of 8 X 8 X 8 mm at full width half maximum (FWHM). Also, the coordinates of the brain produced by SPM were mapped on to the brain using the computer software program Talairach Atlas (http://www.talairach.org/) developed by Jack Lancaster and Peter Fox (Lancaster et al., 1997; Lancaster et al., 2000).

4.8 Preprocessing

Standard preprocessing of functional images was used, which included motion correction, spatial normalization of each individual echo planar imaging (EPI) time series to the MNI (Montreal Neurological Institute) template brain in Talairach orientation, and smoothing with a Gaussian 8 mm full width half maximum (FWHM) filter. In preprocessing no slice-timing correction was applied as is standard in sparse temporal sampling. This is because slice-timing corrections involves interpolation between successive volumes, which would result in error because of the long TR (15s) used. Scaling according to mean voxel value was applied throughout each functional volume to correct for any confounds arising from global vascular effects. The difference in BOLD response between isochronous, double, and LSS conditions, between simple and complex conditions, and between Human-Human, Solo, and Human-Computer conditions were estimated at every voxel across the whole brain volume for each participant with a first level the general linear model (Friston, Frith, Frackowiak, & Turner; 1995a,b), and group data were analyzed with a random effects model at the second level (Friston, Holmes, Price, Büchel, & Worsley, 1999).

4.9 First level analysis

For first level analyses, statistical parametric maps were thresholded by using clusters determined by Z=+2.3 and an uncorrected cluster significance of between $P < 0.05$ and $P < 0.1$ (modulated to produce an analysis with a representative set of clusters in each case). The design matrix was constructed for all 24 runs with effects for the Joint Action (i.e. Human-Human, Solo, Human-Computer) and Complexity (i.e. Simple, Complex). The contrast vectors used to run first level analysis can be found in Appendix B.

4.10 Second level analysis

For second level analyses, statistical parametric maps were thresholded by using clusters determined by Z=+2.3 and a corrected cluster significance of $P < 0.1$. Furthermore, contrast images corresponding to each 'response to outcome' effect on a subject-wise basis were entered into a 2×3 within-subjects t-test (Henson & Penny, 2003) using the outcome (synchronize or no synchronize) and the three levels (Human-Human, Solo, Human-Computer) as independent variables. We constructed a 2×3 within-subjects t-test model at the second level. Thus, taking the average across runs for the 6 'response to outcome' conditions and performing the t-test at the 2nd level was more efficient. This second level model also allowed us to plot the brain activity (beta values) profile across all of the outcome conditions relative to the control conditions.

The second level within-subjects t-test design matrix in SPM (suggested in Henson and Penny, 2005) contained separate columns for the main effects, interactions, constant term, and subject effects. Using this type of model, we constructed a 2×3 within-subjects t-test at the second level with a main effect for joint action and main effect for complexity, constant term, and subject specific terms.

5 Results

5.1 Behavioural Data

5.1.1 Level of Synchronization in Complexity and Joint Action Conditions

We chose to compare the variability in synchronization levels between the Simple and Complex conditions in order to understand how well partners are able to synchronize with each other during an SMS task. In the simple condition each performer must match their onset timing with that of their partner. However, the Complex condition places more timing demands on each performer compared to that of the Simple condition. This is because in the Complex condition each performer must first perceive the onset time of their partner's beat and then mentally compute when their own onset time should occur. The Complex condition therefore demands more coordination between the performer's sensory perceptions and motor executions than the Simple condition.

It was also necessary to assess the variability in synchronization levels within the context of the Joint Action conditions. This is because the levels of interaction between the Human-Human, Solo, and Human-Computer conditions vary. For example, the Human-Human and Human-Computer conditions require participants to be aware of, and monitor their interaction with, their partner. Furthermore, performer's interactions may be different depending on if their partner is a human who can adapt to beat deviations, or is an electronic track, which is unable to adapt to such deviations. Contrastingly, the Solo condition does not require the performer to interact with, and adapt to, a partner's onset timing. Instead, the performer's task is to keep drumming a consistent beat without an external cue.

Because playing in synchrony (Simple condition) requires less cognitive effort than playing in asynchrony (Complex condition) it is hypothesized that the synchronization variability will be less for the Simple condition than the Complex condition. Also, because human partners (Human-Human condition) are able to mutually adapt to beat deviations is it hypothesized that the synchronization variability will be less for the Human-Human condition compared to the Solo and Human-Computer conditions.

The onset timings of both participants were compared in order to assess the accuracy of synchronization within all five conditions (1. Human-Human Simple, 2. Human-Human Complex, 3. Solo Simple, 4. Human-Computer Complex, and 5. Human-Computer Complex). First, the audio tracks of each participant recorded during the fMRI study were transformed into wave files using the audio editor program Audactiy (http://audacity.sourceforge.net/). Using this same program the wave files of each pair of participants were cut at the exact same time. This allowed for the comparison of onset times for each participant pair, and in turn allowed for the measurements of variability in synchronization levels to be calculated for all five conditions. The means and standard deviations for each participant and for each condition were extracted using the MIR toolbox (Eerola & Toivianinen, 2004; Lartillot & Toiviainen, 2007), which functions within the computer program Matlab. The means and standard deviations for the Complexity and Joint Action conditions were then compared using the statistical software Jump (http://www.jmp.com/software/).

Findings show that the Complex conditions were significantly lower in synchronization variability than the Simple conditions. Using error bars in Figure 9 shows that when comparing the Human-Human conditions to the Human-Computer conditions the Complex conditions were significantly lower in synchronization variability compared to that of the Simple conditions. Therefore, both the Human-Computer Complex condition and the Human-Human Complex condition are significantly lower in synchronization variability compared to the Human-Human Simple and Human-Computer Simple conditions. It is important to note that the Solo Simple condition was ignored in this comparison as it lacks a Complex condition.

Furthermore, no significant difference was found in synchronization variability between the two Joint Action Complex conditions or between the two Joint Action Simple conditions when comparing the Human-Human and Human-Computer conditions. However, the Solo Simple condition was found to have significantly lower synchronization variability compared to both the Human-Human Simple condition and the Human-Computer Simple condition. These results are also illustrated in Figure 9.

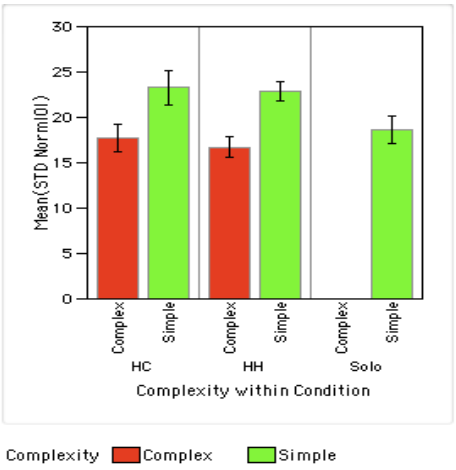


Figure 9: shows the interaction effects of the Complexity conditions and the Joint Action conditions

To make ensure that our synchronization variability findings were accurate we measured how consistent participants were at keeping each rhythmic duration the same length. For example, how consistent were participants in keeping the length of each eighth note the same duration throughout each condition? In order to answer this question the inter-onset interval (IOI) of each rhythm produced was recorded. The mean IOI for each rhythmic duration was then calculated. Results show that participants were extremely consist in their IOI productions for each rhythmic duration. In fact, there were no significant difference in length of the mean IOI between any of the five conditions. These results are shown in Figure 10.

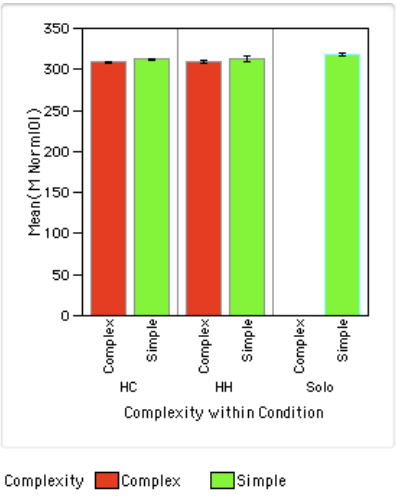


Figure 10: shows the mean length of IOI for the Complexity and Joint Action Conditions

5.1.2 Levels of Synchronization for Isochronous, Double, and LSS conditions

Because we found significant differences in the synchronization variability between the Complexity and Joint Action conditions further investigation of how the different rhythms influenced the synchronization variability were examined. It was hypothesized that the isochronous rhythm would have the lowest synchronization variability, that the LSS rhythm would have the highest synchronization variability, and that the double rhythm would show a synchronization variability somewhere between

the isochronous and LSS rhythms. Using the same method as before we compared the interaction effects of the Complexity, Joint Action, and Rhythm conditions on synchronization variability. Findings show that there were no significant differences in synchronization variability between the Isochronous conditions and the LSS condition. However, there were four significant differences in synchronization variability between the Isochronous conditions and the Double conditions. First, the Isochronous Human-Computer Complex condition was significantly lower in synchronization variability compared to the corresponding Double condition (Double Human-Computer Complex condition). Second, The Isochronous Human-Computer Simple condition was significantly lower in synchronization variability compared to the corresponding Double condition. Third, the Isochronous Human-Human Simple condition was significantly lower in synchronization variability compared to the corresponding Double condition. And fourth, the Isochronous Solo Simple condition was significantly lower in synchronization variability compared to the corresponding Double condition.

Findings also showed three significant differences between the LSS conditions and Double conditions. First, the LSS Human-Computer Complex condition was significantly lower in synchronization variability compared to the corresponding Double condition (Double Human-Computer Complex condition). Second, the LSS Human-Computer Simple condition was significantly lower in synchronization variability compared to the corresponding Double condition. And third, the LSS Human-Human Simple Condition was significantly lower in synchronization variability compared to the corresponding Double condition. These results, which concern the individual rhythmic conditions in synchronization variability, are straightforward, but tedious. Therefore, these results are best understood through Figure 11.

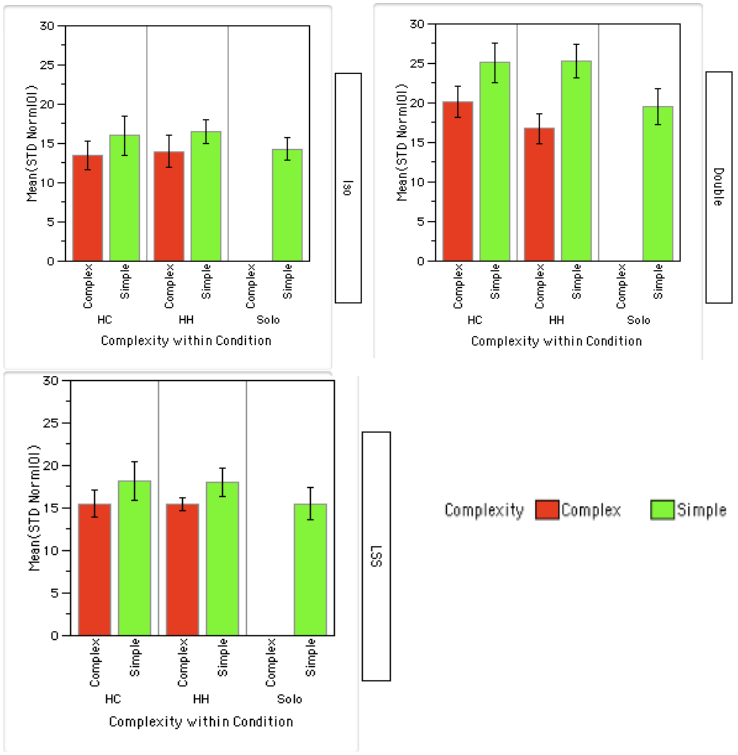


Figure 11: illustrates the variability in synchronization levels for the three different rhythms, isochronous, double, and LSS.

To make ensure that our synchronization variability findings were accurate for the Rhythm condition we again measured how consistent participants were at keeping each rhythmic duration the same length. The previous method used to calculate the mean IOI for each rhythmic duration. Again, results show that participants were extremely consist in their IOI productions for each rhythmic duration.

However, there was one significant difference found between the Double Human-Computer Complex condition and the Double Human-Human Complex condition. These results are shown in Figure 12.

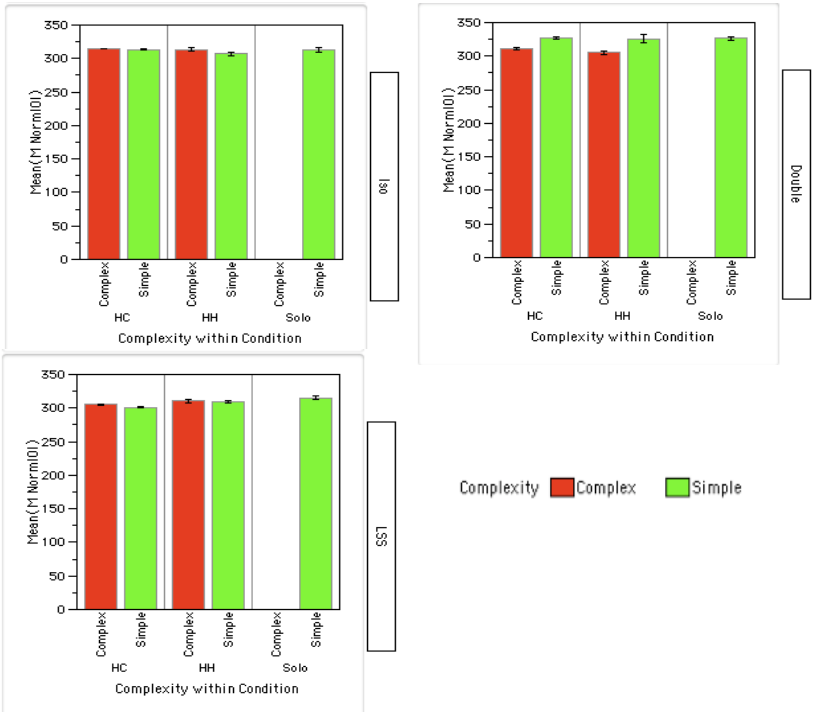


Figure 12: 12 shows the mean length of IOI for the Complexity, Joint Action, and Rhythmc Conditions

5.2 Brain Imaging Data

5.2.1 Simple Verse Complex Condition

Because there were significant differences found within the behavioral data concerning the Simple and Complex conditions we decided to investigate if any significant differences in neural activation could be found in this condition. Because there are more coordination, planning, and timing demands present in the Complex condition it was hypothesized that the Complex condition would require more cognitive resources compared to the Simple condition. More precisely, the cerebellum was expected to be more active during the complex condition because of the increase in timing demands. We also expected increased activation in the SMC and PMC as these areas are involved in the planning and executing of motor movements.

We used SPM to analyze the functional brain data. Screenshots of the exact parameters and procedures performed during pre-processing, 1st level analysis, and 2nd level analysis can be found in Appendix D and Appendix E. When we subtracted the functional brain data of the Simple condition from the Complex condition we found no significant differences in neural activation. Surprisingly, no significant activation was found in the cerebellum or motor cortex, the two areas hypothesized to be involved in motor executions.

5.2.2 Overall Activation of Joint Action

Because of the differences in interaction levels in the Joint Action condition, and more so because there were significant differences found in the Joint Action conditions of the behavioral data, we analyzed the brain imaging data in order to potentially locate any significant differences in neural activation within this condition. We were first interested in examining if the neural activity differed between these three conditions. First an ANOVA was performed in order to determine if any differences in neural activation existed between these three Joint Action conditions. Examining the whole brain we

were able to identify five voxels of statistical significance, which can be found in table 2. Furthermore, the brain areas are illustrated in Figure 13.

Table 2: Coordinates and Z-score values for activations between the Human-Human, Solo, and Human-Computer conditions

Region	BA	x	y	z {mm}	Z-score
Right Cerebrum, Occipital Lobe, Precuneus, Gray Matter	31	6	-64	28	5.63
Left Cerebrum, Parietal Lobe, Postcentral Gyrus, White Matter	*	-12	-50	64	5.38
Left Cerebrum, Occipital Lobe, Lingual Gyrus, Gray Matter	19	-18	-62	0	5.38
Left Cerebrum, Limbic Lobe Posterior Cingulate, White Matter	*	-22	-66	8	5.33
Right Brainstem, Midbrain	*	10	-28	-10	5.12

For all tables Talairach atlas coordinates are in millimeters along the left-right (x), anterior-posterior (y), and superior-inferior (z) axis. The intensity threshold for this table is $Z > 4.27$; $p < 0.05$. * No BA (Broadmann area) defined.

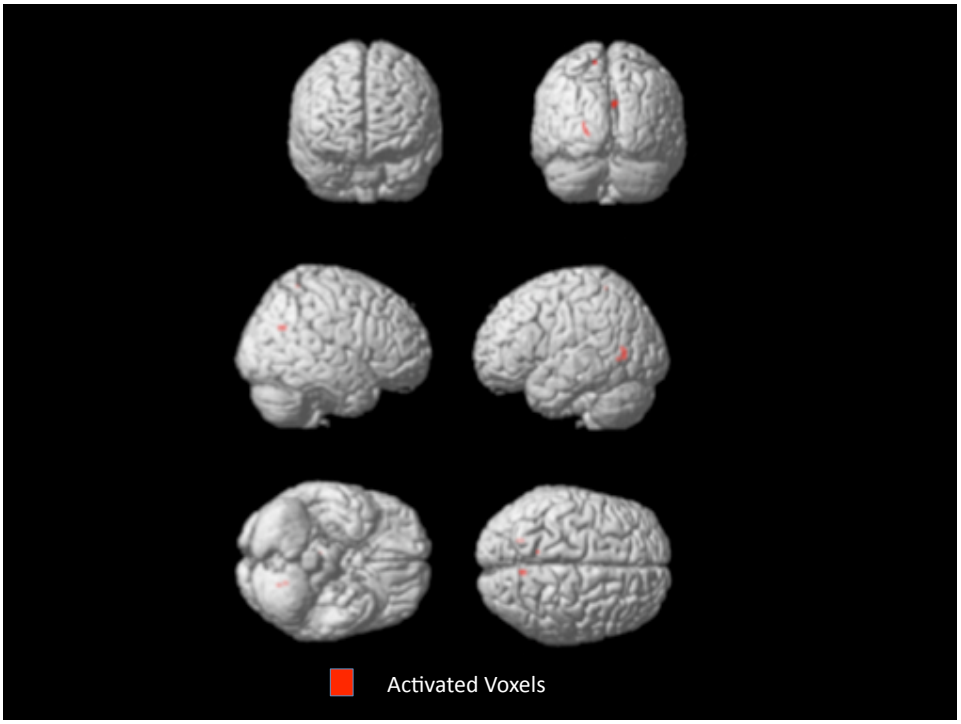


Figure 13: Five voxels activated by the ANOVA comparing the Human-Human condition, the Solo condition, and the Human-Computer condition.

5.2.3 Neural Correlates of the Human-Human Joint Action Condition and Solo Condition

Upon running an ANOVA and finding statistically significance values for the three performance condition we ran t-tests to identify which conditions showed significant activation. We first compared performances of the solo drumming task with a the Human-Human drumming task. In the Solo drumming task a single performer is to coordinate drumming taps with a pre-recorded track. After 15 s the track ceases and the drummer is to continue tapping alone. This tasks requires one to maintain a steady tap by relying on an internal metronome. Contrastingly, in the Human-Human drumming task each performer must perceive the beat of their partner and then adjust their own onset times in order to coordinate their actions with each other. Because of the larger demands placed on performers during the Human-Human drumming task it was hypothesized that the Human-Human condition would require more mental resources than the Solo condition. We particularly expected activation in the posterior parietal lobe, the basal ganglia, and the cerebellum, which are involved in learning and sensorimotor activity. However, when the solo condition was subtracted from the Human-Human condition no significant differences in neural activation were found.

5.2.4 Neural Correlates of the Human-Human Condition and Human-Computer Condition

We were also interested in the differences in neural correlates between the Human-Human condition and the Human-Computer condition. This is because two humans are able to adapt their onset timing to each other, where as a pre-recorded track demands a leader-follower relationship. This means that the human performer must follow by adapting his or her onset times to that of the track. Furthermore, the track is seen as the leader in the performance because although it contains human deviations in onset timing, it is unable to adapt to the timing deviations of its human partner. It was therefore hypothesized that the Human-Human condition would recruit additional neural regions compared to that of Human-Computer condition. Specifically, we expected increased activation in the posterior parietal lobe, basal ganglia, and the cerebellum. However, when the Human-Computer condition was subtracted from the Human-Human condition so significant activation was found.

5.2.5 Neural Correlates of the Solo and Human-Computer Conditions

Lastly, we were interested in the differences in neural activation between the Solo condition and the Human-Computer condition. The major difference between these two conditions is that the Solo condition lacks an external reference and therefore must rely only on internal sources in order to keep time. Therefore, the Solo condition is seen as a more demanding task. We therefore hypothesized that the Solo condition would show increased activation in the putamen, ventrolateral thalamus, SMA, and basal ganglia. However, when we subtracted the Solo condition from the Human-Computer condition, we found no significant difference in activation.

6 Discussion

6.1 Aim

The purpose of this research was to examine the variability in rhythmic synchronization between varying levels of Joint Action (Human-Human, Solo, Human-Computer), varying levels of Complexity (synchronous and asynchronous), while using various Rhythms (isochronous, double, and LSS). The behavioral study was interested in investigating the synchronization variability, and the fMRI study was interested in identifying the neural correlates under these various drumming conditions.

6.2 Behavioral Data

6.2.1 Level of Synchronization in Complexity and Joint Action Conditions

It was hypothesized that the synchronization variability would be less for the Simple condition than the Complex condition. However, the results do not support the hypothesis as findings showed that the Complex condition contained significantly lower synchronization variability than the Simple condition. It is necessary to note that the Solo Simple condition was ignored in this comparison as it lacks a Complex condition. At first glance, this finding suggests that within duet performances it is a cognitively easier task to perform asynchronous rhythms rather than synchronous rhythms. However, Aschersleben & Prinz (1995) previously showed that SMS synchronization also requires coordination between sensory perceptions and motor executions, which suggests that that the timing demands necessary for synchronizing rhythms are not less than that of rhythms performed in asynchrony. Furthermore, it has also been shown that it is more difficult to perform with a partner who is unpredictable and nonresponsive (Kovalinka, Vuust, Roepstorff, & Firth, 2010). This suggests that it is not important whether participants are performing rhythms in synchrony or asynchrony, but that it is important that partners are able to anticipant and adapt to each others beats.

It was also hypothesized that the synchronization variability would be less for the Human-Human condition compared to the Solo condition and the Human-Computer condition. However, this hypothesis was not supported for several reasons. First, there was no significant difference between the Human-Human Simple condition and the Human-Computer Simple condition. Second, there was no significant difference between the Human-Human Complex condition and the Human-Computer Complex condition. These results suggest that when partners perform together, despite whether one partner is human or electronic, synchronization variability is not affected. Lastly, the Solo Simple condition contained synchronization variability that was significantly lower than that of the Human-Human conditions. Although this result does not support the hypothesis, it runs parallel to previous research. When investigating piano performances it was found that participants were better able to synchronize with a recording of themselves compared to synchronizing with another pianist (Knoblich & Repp, 2007) and Lewis and colleagues (2004) have suggested that tapping without an external cue requires only over-learned motor commands. This suggests that is cognitively easier to synchronize with oneself than with a partner.

Findings concerning the mean IOI showed no significant differences between any of the conditions. This suggests that participants were extremely consistent in producing rhythmic durations of the appropriate length. This further suggests that the findings relevant to the synchronization variability were in fact accurate.

6.2.2 Levels of Synchronization for Isochronous, Double, and LSS conditions

Among the three rhythm conditions, Isochronous, Double, and LSS, It was hypothesized that the isochronous rhythm would have the lowest synchronization variability, that the LSS rhythm would have the highest synchronization variability, and that the double rhythm would show a synchronization variability somewhere between the isochronous and LSS rhythms. Results do not support the hypothesis. The results show that the Isochronous condition contained the lowest synchronization variability; however, the Isochronous condition was not significantly lower than the LSS condition. Furthermore, the Double condition was found to have the highest synchronization variability. Previous

research has not compared the differences in synchronization between these rhythmic conditions, nor has previous research compared the interaction effects of Rhythm, Joint Action, and Complexity with each other. However, Dhamala and colleagues (2003) found that additional internal effort is required for increased rates of tapping. This suggests that it is more difficult to maintain tapping consistency in the Double condition because of the short duration which occurs between the first and second tap. Furthermore, Rhythms with ratios of 1:3, which exist in the double rhythm are more difficult for humans to calculate as compared to durations which are multiples of 2. (Repp, London, & Keller, 2008). Therefore, this finding is congruent with previous research because of the increased rate of tapping in the Double condition compared to that of the Isochronous condition and because of the 1:3 ratio of the Double condition, which is more difficult for humans to consistently maintain.

Findings concerning the mean IOI showed only one significant difference between any of the conditions. This significant difference was found between the Double Human-Computer Complex condition and the Double Human-Human Complex condition. This suggests that overall, participants were fairly consistent in producing rhythmic durations of the appropriate length. It further suggests that performing with an unresponsive partner leads to greater variability compared to performing with an adaptable human partner. However, these mean IOI findings still suggest that the synchronization variability results of the individual rhythm conditions are accurate.

6.3 Brain Imaging Data

6.3.1 Neural Correlates of the Simple Condition and Complex Condition

For the brain imaging data concerning the Simple and Complex conditions it was hypothesized that the Complex condition would increase activation in the cerebellum, SMC, and PMC, compared to the Simple condition. Results did not support this hypothesis and instead showed no differences in neural activation between the Simple and Complex condition. Aschersleben & Prinz (1995) have shown that the time demands of synchronizing rhythms and asynchronizing rhythms are not necessarily different. Furthermore, the cerebellum, SMC, and PMC are brain regions associated with increases in planning and executing motor movements (Dhamala et al., 2003). Therefore, these results of the current study suggest that equal amounts of planning and executing are needed for drumming tasks regardless of whether these tasks involve drumming in synchrony or asynchrony.

6.3.2 Neural Correlates of the Human-Human Joint Action Condition and Solo Condition

After performing an ANOVA on the three levels of the Joint Action condition (Human-Human, Solo, and Human-Computer) results showed there to be significant differences between these three levels. Therefore, a t-test was performed in order to identify which levels of the Joint Action condition were significantly different from one another. When comparing the Human-Human condition to the Solo condition it was predicted that the Human-Human condition would activate the posterior parietal lobe, the basal ganglia, and the cerebellum. Although these brain regions have been shown to increase in activation during learning and sensorimotor tasks (Gold & Shadlen, 2007; Hikosaka et al., 2002) and have shown to be directly involved in joint actions tasks (Kial, Gazzola, & Keysers, 2009) no significant differences in these brain regions were found between the Human-Human condition and the Solo condition. This either suggests that the Human-Human condition and the Solo condition recruit similar neural mechanisms for drumming tasks or it suggests that the current study was unable to distinguish any differences in activation between these two condition. Because this specific analysis resulted in no significant differences in any brain regions it implies that the current study failed in some way. The current study may have failed to employ distinct enough parameters to promote differences in neural activation between the Human-Human and Solo conditions. It may have also failed to apply the correct statistical measurements between the two conditions, which resulted in non-significant values for all voxels measured. When this same t-test was performed again using an FWE-corrected p value of 0.2 there were significant differences in neural activation found between the Human-Human condition and the Solo condition. However, the significant voxels found were sparse and therefore interpreted as noise. Although this second t-test does not shed light on why no significant differences were found

between these two conditions it does suggest that the current study needs to be improved upon before further research is conducted.

6.3.3 Neural Correlates of the Human-Human Condition and Human-Computer Condition

Before performing a t-test to examine the differences between the Human-Human condition and the Human-Computer condition it was predicted that the Human-Human condition would again recruit the posterior parietal lobe, basal ganglia, and the cerebellum. This is because the Human-Computer condition is not considered a joint action performance, as the audio track cannot adapt to its human partner. This hypothesis was also not supported as the analysis resulted in no significant differences in neural activation between these two conditions. Performing the same t-test again using an FWE-corrected p value of 0.2 resulted in no significant differences in neural activation between the Human-Human condition and the Human-Computer condition. If both conditions are considered duet performances, in the sense that that there are two performers, then these results suggest that the neural mechanisms involved in both the Human-Human condition and the Human-Computer condition are similar. However, if the Human-Computer condition is interpreted as a leader-follower performance (Kovalinka, Vuust, Roepstorff, & Firth, 2010) where mutual adaptation does not take place, then these results may further suggest that the current study is again, flawed in some way.

6.3.4 Neural Correlates of the Solo and Human-Computer Conditions

When comparing the Solo condition and the Human-Computer condition it was predicted that the Solo condition would show increased neural activation in the putamen, ventrolateral thalamus, SMA, and basal ganglia. However, this hypothesis was not supported as no significant differences were found between these two conditions. This finding may suggest that the neural activation for the Solo condition and the Human-Computer condition are similar, which may be because both conditions lack mutual adaptation. A second t-test was also performed for this analysis using an FWE-corrected p value of 0.2. Again, no significant differences were found between these two condition during this second analysis.

It is difficult to determine the cause of the non-significant differences in neural activation among the Joint Action conditions because although previous research has shown that joint performances do activate several brain regions, including parts of the temporal, auditory, and visual cortex (Kokal, Gazzola, & Keysers, 2009; Newman-Norlund et al., 2008; Newman-Norlund et al., 2007) there is only a small body of literature concerning joint action drumming tasks. Furthermore, the previous research conducted on the neural correlates of joint action tasks has not necessarily focused on musical aspects, and instead has investigated the social consequences of joint action. (For an example see Kokal, Gazzola, & Keysers, 2009). Therefore, subsequent research is much needed in order to better understand the neural correlates of joint action drumming.

6.4 Limitations

Before concluding it is necessary to address a few important limitations. One major limitation of this research is that it did not compare an overall drumming task to rest (baseline) conditions. Few brain volumes were actually acquired for rest conditions. In fact, only one in five volumes acquired were of rest conditions. This means that there was not enough data collected about neural activation during rest conditions to be compared to task conditions.

The conventional sparse sampling method used to acquire brain imaging data inherently contains limitations of its own. For example, the method only provides one measure of the BOLD response for each epoch/ stimulus presentation. Therefore, sparse temporal sampling does not allow for differences in timing profiles of the BOLD response to be assessed unless each stimulus is presented several times (Schwarzbauer, Davis, Rodd, & Johnsrude, 2006). In fact, BOLD changes are only found to be approximately one percent above baseline levels (Ogawa et al., 1990). In this therefore important to repeat measurement of the BOLD signal many times. In the current study repetition of measurements were only taken for the Human-Human conditions. Furthermore, the time course of the BOLD response for different brain regions varies (Josephs & Henson, 1999) means that the acquisition of a brain volume

may not have occurred at the peak of the BOLD response for all brain areas. Therefore, conventional sparse sampling has shown to result in a lack of temporal information because one brain volume is scanned after a signal stimulus presentation. It has been suggested that the presentation of stimuli should be "jittered" with respect to the time at which image acquisition begins (Belin, Zatorre, Hoge, Evans, & Pike, 1999). These limitation of the sparse temporal sampling technique may have contributed to the small amount of significant voxels identified during the analysis process.

Not only does the method used to acquire brain volumes contain limitations, but so does the brain imaging technique of fMRI. First, fMRI only measure the change in blood flow within the brain. It is therefore an indirect measure of neuronal activity and can be influenced by non-neural changes within the body. The BOLD itself can be affected by external factors such as medications, attention and levels, and fluctuations of C02 in the blood stream. Also, it is important to note that fMRI only produces correlational statistics and therefore cannot imply causality of a stimulus-response relationship. The limitation of the fMRI technique may have contributed to the small amount of significance found within the brain imaging data.

Another problem with the fMRI technique is that participants were required to lie supine in the scanner bed while drumming. Although the drumming tasks only required small movements of the wrist, the activity of drumming caused participants' bodies to move and therefore may have caused slight head movement. Head movement can easily lead to reductions in brain imaging quality, which may have further contributed to the lack of significant results found in the current fMRI study.

The participants used in this study pose two problem. First, fMRI experiments are expensive to conduct. Because of this financial limitation only eight participants were recruited. Such a small sample of the population results in a lack of external validity. Second, it could be argued that because the subject pool consisted only of musicians, our findings are of a group that is not representative of the general population.

6.5 Future Research

Future research may expand on the current study in a number of ways. For example, enhanced temporal resolution is offered by a newer interleaved imaging procedure (Schwarzbauer, Davis, Rodd, Johnsrude, 2006). Furthermore, a newer clustered version of sparse sampling is known to give better statistical results compared to conventional sparse sampling (Zaehle et al., 2007). Future research is therefore encouraged to employ the latest developments in sparse temporal sampling as they are likely to produce more accurate and more significant findings compared to the current study.

Because few significant results were found using an fMRI method, future studies may wish to use EEG in conjunction with fMRI. EEG studies are noninvasive and less expensive than fMRI studies, making them a likely technique to be employed in future research. Furthermore, EEG is a brain imaging technique know to have high temporal resolution, while fMRI is known to have high spatial resolution. Therefore, using both imaging techniques future studies may find more accurate results. For example, using both EEG and fMRI together may result in converging evidence about the neural correlates of joint action drumming and therefore result in more conclusive findings.

Comparing the neural correlates of different rhythmic conditions is also highly suggested for future research. This is because the Isochronous, Double, and LSS rhythms required varying amounts of movement and therefore may have recruited different neural substrates. For example, the Double condition required more movement than the Isochronous condition. This turn may have caused the Double condition to be prone to more synchronization variability. For example, there may have be more instances of period corrections or phase corrections. Because previous research has intensively investigated the behavioral aspects of period and phase correction it seems likely that future neuroscientific research would benefit from build upon the existing research.

Future research is also encouraged to use a larger and more broad sample of participants. Data acquired from only eighth musicians is not enough to assume generalizability. And although the current study found it advantageous to use musicians because these participants could quickly learn the rhythms of the drumming task, advanced knowledge of music theory and music performance are not common among the general population. Therefore, using a wider base of participants who have an average background in music may provide more significant behavioral and neural results which are

more generalizable to the human population.

6.6 Conclusion

Humans possess a great ability to coordination actions with others. One important purpose of joint action in humans is to facilitate social, and more broadly, cultural interaction in humans, for example, in producing musical performances. This study therefore aimed to better understand the behavioral and neural aspects involved in joint performances. In order to achieve this aim participants were asked to perform drumming tasks under various conditions. The outcome of these tasks yielded surprising results, which supported none of the purposed hypotheses. This made for difficult interpretations and left many unanswered questions. For example, are the neural correlates of joint action in everyday life reflected in musical ensemble performances? And what are the evolutionary reasons for why humans are able to synchronize so easily with each other? Because joint action and music pervade everyday life it is to important to gain answers to the questions. This suggests that future research should not view the current study as a failed attempt at gaining insight into the coordinated tasks of musical performance. Instead, future research should continue asking questions concerning the relationship between music and joint action, and should do so by learning from the current studies short comings.

Supplementary Information

ANOVA	analysis of variance
BOLD	blood oxygenation level dependent
CBF	cerebral blood flow
CBV	cerebral blood volume
CT	computer tomography
EEG	electroencephalography
EPI	echo planar imaging
fMRI	functional magnetic resonance imaging
FWE	family-wise error rate
FWHM	full-width-at-half-maximum
GLM	general linear model
HRF	haemodynamic response function
IFG	inferior frontal gyrus
IPL	inferior parietal lobe
MEG	magnetoencepholograhy
MNI	montreal neurological institute
MTG	middle temporal gyrus
PAC	primary auditory cortex
PET	position emission tomography
PMC	primary motor cortex
S1	primary somatosensory cortex
SMA	supplementary motor area
SMC	sensorimotor cortex
SNR	signal to noise ratio
SPM	statistical parametric mapping
STG	superior temporal gyrus
TA	time to acquisition
TE	time to echo
TR	time to repetition

7 Bibliography

Alexander, G., DeLong, M., & Strick, P. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annual Review of Neuroscience, 9, 357-381.

Aschersleben, G. & Prinz, W. (1995). Synchronizing actions with events: The role of sensory information. Perception and Psychophysics, 57(3), 305-317.

Bandettini, P., Jesmanowicz, A., Van Kylen J., Birn, R. M., Hyde, J. S. (1998). Functional MRI of brain activation induced by scanner acoustic noise. Magnetic Resonance in Medicine, 39(3), 410-416.

Belliveau, J. W., Knwong, K. K., Kennedy, D. N., Baker, R. R., Stern C. E., Benson, R., Chesler, D. A., Weisskoff, R. M., Cohen, M. S., Tootell, R. B. H., Fox, P. T., Brady T. J., & Rosen, B. R. (1992). Magnetic-Resonance-Imaging Mapping of Brain- Function – Human Visual-Cortex. Investigative Radiology, 27, S59-S65.

Bénar, C., Bagshaw, A. P., & Lemieux, L. (2010). Experimental design and data analysis strategies. In L. Lemieux & C. Mulert (Eds.), EEG-fMRI Physiological Basis, Technique and Applications (221-258). London: Springer.

Belin, P., Zatorre, R. J., Hoge R., Evans A. C., Pike, B. (1999). Event-related fMRI of the auditory cortex. Neuroimage, 10(4), 417-429.

Bernieri, F. J. (1988). Coordinated movement and rapport in teacher-student interactions. Journal of Nonverbal Behavior, 12(2), 120-138.

Bloedel, J. R. (1992). Functional heterogeneity with structural homogeneity: How does the cerebellum operate? Behavioral and Brain Sciences, 15, 666-678.

Brooks, D., Ibañez, V., Sawle, G., Quinn, N., Lees, A., Mathias, C., Bannister, R., Marsden, C., & Frackowiak, R. (1990). Differing patterns of striatal F-dopa uptake in Parkinson's disease, multiple system atrophy and progressive supranuclear palsy. Annals of Neurology, 28, 547-555.

Brown, S. (2000). The "Musilanguage" Model of Music Evolution. In N. L. Wallin, B. Merker, & S. Brown (Eds.), The Origin of Music (271-300). London: The MIT Press.

Brown, R. G., Marsden, C. D. (1990). Cognitive function in Parkinson's disease: from description to theory. Trends in Neuroscience, 13, 21-29.

Byrne, R. W., & Russon, A. E. (1998). Learning by imitation: A hierarchical approach. Behavioral & Brain Sciences, 21, 667-721.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L, Fogassi, L., Gallese, V., Seitz, R., Zilles, K., Rizzolatti, G., & Freundy, H. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. European Journal of Neuroscience, 13(2), 400-404.

Casini, L. & Ivry, R. B. (1999). Effects of divided attention on temporal processing in patients with lesions of the cerebellum or frontal lobe. Neuropsychology, 13, 10-21.

Cho. Z., Chung, S., Lim, D., Wong, E. (1998). Effects of the acoustic noise of the gradient systems on fMRI: A study on auditory, motor, and visual cortices. Magnetic Resonance in Medicine 39(2), 331-335.

Clarke, E., Dibben, N. & Pitts, S. (2010). Music and Mind in Everyday Life. New York: Oxford University Press.

Desmurget, M., Epstein, C. M., Turner, R. S., Prablanc, C., Alexander, G. E., & Grafton, S. T. (1999). Role of the posterior parietal cortex in updating reaching movements to a visual target. Nature Neuroscience, 2, 563-567.

Dhamala, M., Pagnoni, G., Wiensenfeld, K., Zink, C. F., Martin, M., & Berns, G. S. (2003). Neural correlates of the complexity of rhythmic finger tapping. NeuroImage, 20, 918- 926.

Edmister, W. B., Talavage, T. M., Ledden, P. J., Weisskoff, R. M. (1999). Improved auditory cortex imaging using clustered volume acquisitions. Human Brain Mapping, 7(2), 89-97.

Eerola, T. & Toivianinen, P. (2004). MIR in Matlab: The Midi Toolbox. Proceddings of 5th International Conference on Music Information Retrieval, 22-27.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. Journal of Neurophysiology, 73(6), 2608.

Fitch, W. T. (2009). Biology of Music: Another One Bites the Dust. Current Biology, 19(10), R403-4.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: From action organization to intention understanding. Science, 308, 662- 667.

Friston, K., Holmes, A., Price, C., Büchel, C., & Worsley, K. (1999). Multi-subject fMRI studies of conjunction analyses. NeuroImage, 10, 385-396.

Friston, K., Holmes, A. Worsley, J., Poline, J. Firth, C., & Frackowiak, R. (1994). Statistical parametric maps in functional imaging: Ageneral linear approach. Human Brain Mapping, (4), 189-210.

Friston, K., Frith, C., Frackowiak, R., & Turner, R. (1995a). Characterizing evoked hemodynamics with fMRI. NeuroImage, 2, 157-165.

Friston, K., Frith, C., Frackowiak, R., & Turner, R. (1995a). Characterizing dynamic brain responses with fMRI. NeuroImage, 2, 166-72.

Gallese, V., Fadiga, L., Fogassi, L., Rizzolatti, G. (1996). Action recognition in the premotor cortex. Brain, 119(2), 593-609.

Gerloff, C., Corwell, B., Chen, R., Hallett, M., & Cohen, L. G. (1997). Stimulation over the human supplementary motor area interferes with the organization of future elements in complex motor sequences. Brain, 120, 1580-1602.

Goebl, W. & Palmer, C. (2009). Synchronization of timing and motion among performing musicians. Music Perception, 26(5), 427-438.

Gold, J. & Shadlen, M. (2007). The neural basis of decision making. Annual Review of Neuroscience, 30, 535-574.

Grafton, S., Arbib, M., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. 2. Observation compared with imagination. Experimental Brain Research, 112(1), 103-111.

Hall, A. D., Haggard, M. P., Akeroyd, M. A, Palmer, A. R., Summerfield, A. Q., Elliott, M. R., Gurney, E. M., & Bowtell, R. W. (1999). "Sparse" temporal sampling in auditory fMRI. Human Brain Mapping, 7, 213-223.

Hall, A. D., Summerfield, A., Gonçalves M. S., Foster, J., Palmer, A. R., Bowtell, R. W. (2000). Time-course of the auditory BOLD response to scanner noise. Magnetic Resonance in Medicine, 43(4), 601-606.

Hallett, M. (1994). Movement-related cortical potentials. Electromyography and Clinical Neurophysiology, 34, 5-13.

Halsband, U., Ito, N., Tanji, J., & Freund, H. J. (1993). The role of premotor cortex and the supplementary motor area in the temporal control of movement in man. Brain, 116, 243-266.

Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. Proceeds of the National Academy of Sciences USA, 95(5), 15061-15065.

Harrington, D., Rao, S., Haaland, K., Bobholz, J., Mayer, A., & Binderx, J., & Cox, R. (2000). Specialized neural systems underlying representations of sequential movements. Journal of Cognitive Neuroscience, 12(1), 56-77.

Haslinger, B., Erhard, P., Weilke, F., Ceballos-Baumann, A. O., Bartenstein, P., & Grafin, v. E. Schwaiger, M., Conrad, B., & Boecker, H. (2002). The role of lateral premotor- cerebellar-parietal circuits in motor sequences: A parametric fMRI study. Brain Research and Cognitive Brain Research, 13, 159- 168.

Henson, R. & Penny, w. (2003). ANOVAs and SPM. Technical Report, Wellcome Department of Imaging Neuroscience.

Hermann, C. S., Brechmann, A., & Scheich, H. (2010). Simultaneous EEG and fMRI of the Human. In L. Lemieux & C. Mulert (Eds.), EEG-fMRI Physiological Basis, Technique and Applications (173-200). London: Springer.

Hikosaka, O., Nakamura, K., Sakai, K., & Nakahara, H. (2002). Central mechanisms of motor skill learning. Current Opinion in Neurobiology, 12(2), 217-222.

Himberg, T. & Thompson, M. (2009). Group synchronization of coordinated movements in a cross-cultural choir workshop. In 7th Triennial Conference of the European Society for the Cognitive Sciences of Music, (ESCOM). Jyväskylä, Finland. Spoken presentation.

Hunter, M. D., Eickhoff, S. B., Miller, T. W. R., Farrow, T. F. D., Wilkinson, I. D., & Woodruff, P. W. R. (2006). Neural activit in speech-sensitive auditory cortex during silence. Proceedings of the National Academy of Sciences, 103(1), 189-194.

Iacoboni, M., Woods, R. P., Brass, M., Kekkering, H., Mazziotta, J., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. Science, 286, 2526-2528.

Isabella, R. A., Belsky, J., & van Eye, A. (1989). Origins of mother-infant attachment: An examination of interactional synchrony during the infant's first year. Developmental Psychology, 25, 12-21.

Ivry, R. & Keele, S. (1989). Timing functions of the cerebellum. Journal of Cognitive Neuroscience, 1, 136-152.

Ivry, R., Keele, S., & Deiner, H. (1988). Dissociation of the lateral and medial cerebellum in movement timing and movement execution. Experimental brain Research, 73, 167- 180.

Josephs, O. & Henson, R. N. A. (1999). Event-related functional magnetic resonance imaging: Modelling, inference, and optimization. Philosophical Transactions of the Royal Society B: Biological Sciences, 354, 1215-1228.

Jueptner, M., Frith, C., Brooks, D., Frackowiak, R., & Passingham, R. (1997a). Anatomy of motor learning. II. Subcortical structures and learning by trial and error. Journal of Neurophysiology, 77, 1325-1337.

Jueptner, M., Frith, C., Brooks, D., Frackowiak, R., & Passingham, R. (1997b). Anatomy of motor learning. I. Frontal cortex and attention to action. Journal of Neurophysiology, 77, 1313-1324.

Jungers, M. K., Palmer, C., & Speer, S. R. (2002). Time after time: The coordinating influence of tempo in music and speech. Cognitive Procesing, 1-2, 21-35.

Keller, P. (1999). Attending in complex musical interactions: The adaptive dual role of meter. Australian Journal of Psychology, 51(3), 166-175.

Keller, P. E. (2008). Joint action in music performance. In F. Morganti, A. Carassa, & G. Riva (Eds.), Enacting Intersubjectivity: A cognitive and social perspective to the study of interactions (pp. 205-221). Amsterdam: IOS Press.

Keller, P. E., Knoblich, G., & Repp, B. H. (2007). Pianists duet better when they play with themselves: On the possible role of action simulation in synchronization. Consciousness and Cognition 16, 102-111.

Keysers, C., Wicker, B., Gazzola, V., Anton, J. L., Fogassi, L., & Gallese, V., 2004. A touching sight: SII/PV activation during the observation and experience of touch. Neuron, 43, 335-346.

Kim, S. G., Ugurbil, K., Strick, P. L. (1994). Activation of a cerebellar output nucleus during cognitive processing. Science, 265, 949-951.

Kirschner, S. & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. Journal of Experimental Child Psychology, 102, 299-314.

Kokal, I., Engle, A., Kirshner, S., & Keysers, C. (submitted). Synchronized drumming enhances activity in the caudate and facilitates prosocial commitment- If the rhythm comes easy.

Kokal, I., Gazzola, V., & Keysers, C. (2009). Acting together in and beyond the mirror neuron system. Neuroimage, 47(4), 2046-2056.

Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me: Continuous mutual prediction and adaptation in joint tapping. The Quarterly Journal of Experimental Psychology, 63(11), 2220-2230.

Lancaster, J. L., Rainey, L. H., Summerlin, J. L., Freitas, C. S., Fox, P. T., Evans, A. C., Toga, A. W., Mazziotta, J. C. (1997). Automated labeling of the human brain: A preliminary report on the development and evaluation of a forward-transform method. Human Brain Mapping, 5, 238-242.

Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., Kochunov, P. V., Nickerson, D., Mikiten, S. A., Fox, P. T. (2000). Automated Talairach Atlas lables for functional brain mapping. Human Brain Mapping, 10, 120- 131.

Large, E. W. & Jones, M. R. (1999). The dynamics of attending: How people track time- varying events. Psychological Review, 106(1), 119-159.

Lartillot, O. & Toiviainen, P. (2007). A Matlab toolbox for musical feature extraction from audio. Porceedings of the Inernational Conference on Digital Audio Effects.

Leiner, H., Leiner, A., & Dow, R. (1995). The underestimated cerebellum. Human Brain Mapping, 2, 244-254.

Lewis. P. A., Wing, A.M., Pope, P. A., Praamstra, P., & Miall, R. C. (2004). Brain activity correlates differentially with increasing temporal complexity of rhythms during initialisation, synchronisation, and continuation phases of paced finger tapping. Neuropsychologia, 42, 1301-1312.

Lund, T., Norgaard, M. Rostrup, E., Rowe J., Paulson, O. (2005). Motion or activity: Their role in intra- and inter-subject variation in fMRI. Neuroimage, 26, 960-964.

Mann, N. I., Dingess, K., & Slater, P. J. (2006). Antiphonal four-part synchronized chorusing in a Neotropical wren. Biology Letters, 2(1), 1-4.

Mansfield, P., Glover, P. M., Beaumont, J. (1998). Sound generation in gradient oil structures for MRI. Magnetic Resonance in Medicine, 39(4), 539-550.

Mayville,. M., Jantzen, K. J., Fuchs, A., Steinberg, F. L., & Kelso, J. A. S. (2002). Cortical and subcortical networks underlying syncopated and synchronized coordination revealed using fMRI. Human Brain Mapping, 17, 214-229.

Merker, B. H., Madison, G., & Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. Cortex, 45, 4-17.

Middleton, F. A. & Strick, P. L. (2000). Basal ganglia and cerebellar loops: Motor and cognitive circuits. Brain Research Reviews, 31, 236-250.

Molinari, M., Leggio, M. G., & Thaut, M. H. (2007). The cerebellum and neural networks for rhythmic sensorimotor synchronization in the human brain. The Cerebellum, 6, 18-23.

Müller, K., Schmitz, F., Schnitzler, A., Freund, H., Aschersleben, G., & Prinz, W. (2000). Neuromagnetic correlates fo sensorimotor synchronization. Journal of Cognitive Neuroscience, 12(4), 546-555.

Newman-Norlund, R., Bosga, J., Meulenbroek, R. & Bekkering, H. 2008. Anatomical substrates of cooperative joint-action in a continuous motor task: Virtual lifting and balancing. NeuroImage, 41, 169-177.

Newman-Norlund, R., Noordzij, M., Meulenbroek, R., & Bekkering, H. (2007a). Exploring the brain basis of joint action: Co-ordination of actions, goals and intentions. Social Neuroscience, 2(1), 48-65.

Newman-Norlund, R., van Schie, H., van Zuijlen, A., & Bekkering, H. (2007b). The mirror neuron system is more active during complementary compared with imitative action. Nature Neuroscience, 10(7), 817-818.

O'Boyle, D. J., Freeman, J. S., Cody, F. W. J. (1996). The accuracy and precision of timing of self-paced, repetitive movements in subjects with Parkinson's disease. Brain, 119, 51-70.

O'Conner, K. N. & Sutter, M. L. (2000). Global spectral and location effects in auditory perceptual grouping. Journal of Cognitive Neuroscience, 12, 342-354.

Okada, T. & Nakai, T. (2003). Slient fMRI acquisition methods for large acoustic noise during scan. Magnetic Resonance in Medical Sciences, 2(4), 181-187.

Patel, A., Iversen, J., Bregman, M., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. Current Biology, 19, 827- 830.

Pellegrino, G., Fadiga, >, Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. Experimental brain research, 91(1), 176- 180.

Peretz, I. (1990). Processing of local and global musical information in unilateral brain damaged patients. Brain, 13, 1185-1205.

Petrini, K., Dahl, S., Rocchesso, D., Waadeland, C. H., Avanzini, F. A., Puce, A., & Pollick, F. E. (2009). Multisensory integration of drumming actions: Musical expertise affects perceived audiovisual asynchrony. Experimental Brain Research, 198, 339-352.

Phillips-Silver, J., Aktpis, C. A., & Bryant, G. A. (2010). The ecology of entrainment: Foundations of coordinated rhythmic movement. Music Perception, 28(1), 3-14.

Pop, D. & Martin, M. (1997). Everybody (Backstreet's Back) [Recorded by Backstreet Boys]. On Backstreet boys [CD]. Orlando, Florida, USA: Jive Records.

Praamstra, P., Turgeon, M., Hesse, C., Wing, A., & Perryer, L. (2003). Neurophysiological correlates of error correction in sensorimotor synchronization. NeuroImage, 20, 1283- 1297.

Price, D. L., De Wilde, J. P., Papadaki, A. M., Curran J. S., Kitney R. I. (2001). Investigation of acoustic noise on 15 MRI scanners from 0.2T to 3 T. Journal of Magnetic Resonance Imaging, 13(2), 288-293.

Puce, A. & Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 358, 435-445.

Ramnani, N. & Passingham, R. E. (2001). Changes in the human brain during rhythm learning. Journal of Cognitive Neuroscience, 13, 952-966.

Rao, S. M., Binder, J. R., Bandettini, P. A., Hammeke, T. A., Yetkin, F. Z., Jesmanowics, A., Lisk, L. M., Morris, G. L., Mueller, W. M., Estkowski, L. D., Wong, E. C., Haughton, V. M., & Hyde, J. S. (1993). Functional magnetic resonance imaging of complex human movements. Neurology, 43, 2311-2318.

Rao, S. M., Harrington, D. L., Haaland, K. Y., Bobholz, J. A., Cox, R. W., & Binder, J. R. (1997). Distributed neural systems underlying the timing of movements. The Journal of Neuroscience, 17(14), 5528-5535.

Repp, B. (2001). Phase correction, phase resetting, and phase shifts after subliminal timing perturbations in sensorimotor synchronization. Journal of Experimental Psychology: Human Perception and Performance, 27(3), 600-621.

Repp, B. (2005). Sensorimotor synchronization: A review of the tapping literature. Psychological Bulletin & Review, 12(6), 969-992.

Repp, B. & Keller, P. (2004). Adaptation to tempo changes in sensorimotor synchronization: Effects of intention, attention, and awareness. Quarterly Journal of Experimental Psychology, A, 57, 499-521.

Repp, B. & Keller, P. (2008). Sensorimotor synchronization with adaptively timed sequences. Human Movement Science, 27, 423-456.

Repp, B., London, J., & Keller, P. (2008). Phase correction in sensorimotor synchronization with nonisochronous sequenze. Music Perception, 26(2), 171-175.

Repp, B., & Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms, Psychological Research, 68, 252-270.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., & Fazio, F. (1996). Localization of grasp representations in humans by PET: Observation versus execution. Experimental Brain Research, 111, 246-252.

Rizzolatti, G. & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. Nature Reviews Neuroscience, 11(4), 264-274.

Schwarzbauer, C., Davis, M. H., Rodd, J. M., & Johnsrude, I. (2006). Interleaved silent steady state (ISSS) imaging: A new sparse imaging method applied to auditory fMRI. NeuroImage, 29(3), 774-782.

Sebanz, N., Rebbechi, D., Knoblich, G., Prinz, W., & Frith, C. D. (2007). Is it really my turn? An event related fMRI study of task sharing. Social Neuroscience, 2, 81-95.

Sadato, N., Campbell, G., Ibanez, V., Deiber, M., & Hallett, M. (1996). Complexity affects regional cerebral blood flow change during sequential finger movements. Journal of Neuroscience, 16, 2691-7000.

Shah, N. J., Jäncke, L., Grosse-Ruyken M., Müller-Gärtner, H. (2000). Influence of acoustic masking noise in fMRI of the auditory cortex during phonetic discrimination. Journal of Magnetic Resonance Imaging, 9(1), 19-25.

Smith, A. P. (1988). Acute effects of noise exposure: An experimental investigation of the effects of noise and task parameters on cognitive vigilance tasks. International Archives of Occupational and Environmental Health,60, 307-310.

Strick, P., Hoover, J., Mushiake, H. (1993). Evidence for "output channels" in the basal ganglia and cerebellum. In: Role of the cerebellum and basal ganglia in voluntary movement (Mano, N., Hamada, I., & DeLong, M. R., (Eds.), pp. 171-180. New York: Elsevier Science.

Thach, W. T., Goodkin, H. & Keating, J. (1992). The cerebellum and the adaptive coordination of movement. Annual Review of Neuroscience, 15, 403-442.

Thaut, M. H., Stephan, K. M.. Wunderlich, G., Schicks, W., Tellmann, L., Herzog, H., McIntosh, G., Seitz, R. J., & Homber, V. (2009). Distinct cortico-cerebellar activations in rhythmic auditory motor synchronization. Cortex, 45, 44-53.

Thioux, M., Gazzola, V., & Keysers, C. (2008). Ation understanding: How, what and why. Current Biology, 18(10), R431-434.

Tomasello, M., Carpenter, M., Call, J. Behne, T., Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. Behavioral and Brain Sciences, 28(5), 675-735.

Turner, R., Grafton, S., Votaw, J., Delong, M., & Hoffman, J. (1998). Motor subcircuits mediating the control of movement velocity: A PET study. Journal of Neurophysiology, 80, 2162-2176.

Umilta, M., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing: A neurophysiological study. Neuron, 31(1), 155- 156.

Ullen, F. Forssberg, H., & Ehrsson, H. H. (2002). Neural networks for the coordination of the hands in time. Jounral of Neurophysiology, 89, 1126-1135.

Wilkinson, I., Hadjivassiliou, M., Dickson, J., Grünewald, R., Wallis, L. & Griffiths, P. (2005). Cerebellar abnormalities on proton MR spectroscopy in gluten ataxia. Journal of neurology, Neurosurgery, and Psychiatry, 76, 1011-1013.

Wilson, M. (2001). Perceiving imitatible stimuli: Consequences of isomorphism between input and output. Psychological Bulletin, 127(4), 543-553.

Worsley, K, Evans, A., Marrett, S., Neelin, P. (1992). A three-dimensional statistical analysis for CBF activation studies in human brain. Journal of Cerebral Blood Flow and Metabolism, 12, 900-918.

Zaehle, T. Schmidt, C. F., Meyer, M., Baumann, S., Baltes, C., Boesiger, P., Jancke, L. (2007). Comparison of "silent clustered and sparse temporal fMRI acquisitions in tonal and speech perception tasks. Neuroimage, 37(4), 1195-1204.

Zaehle, T., Wüstenberg, T., Meyer, M., Jäncke, L. (2004). Evidence for rapid auditory perception as the foundation of speech processing: A sparse temporal sampling fMRI study. European Journal of Neuroscience, 20, 2447-2456. Zhang, Y. (2011, March 8). General Linear Model [Powerpoint slides]. Retrieved from http://vista.shef.ac.uk/webct/urw/lc19897.tp0/cobaltMainFrame.dowebct

APPENDIX A

Participant information sheet

Instructions for Joint Drumming Behavioral Pilots

Inter-personal Coordination of Rhythmic Action in Music

You are being invited to take part in a research project that investigates how people coordinate the timing of their actions when they are making music together, as compared to playing alone or with a computer-based partner. By examining the timing of drumming sounds made by pairs of people when following and playing different rhythm patterns, we can characterise the processes involved. Please read the following information carefully, and ask us if there is anything that is not clear or if you would like more information.

We have chosen to invite you because you have a background in music performance.

If you do decide to take part you will be given this information sheet to keep (and be asked to sign a consent form) and you can still withdraw at any time without it affecting any benefits that you are entitled to in any way. You do not have to give a reason. Prof Parsons and Dr Timmers are the contact persons to join the study and to receive further information. You can also contact them for raising complaints concerning the research procedure or concerning your well-being related to the study. If you feel your complaints have not been dealt with appropriately you could contact the University's 'Registrar and Secretary'.

The study will take place in a single, one-hour session at the University of Sheffield. You will be asked to listen to sequences of sounds and tap on a drum with your hand(s), often along with another person (your performance partner). Your performances will be recorded for later analyses.

We will ask you to provide some information about your musical background and age. This information will assist with analysing our findings. All information will be kept strictly anonymous and confidential. You will not be able to be identified in any reports or publications.

There are no disadvantages or risks in taking apart in our study. Likewise, there are no benefits to participation. Your participation will help scientific researchers such as ourselves gain and disseminate new insight into joint rhythmic action. The findings may have implications in musical performance and education, and clinical medical contexts.

If you have any concerns or complaints about your participation you may discuss them with Professor Lawrence Parsons (Department of Psychology) or Dr. Renee Timmers (Department of Music). If you require further consultation, you may discuss them with the University Registrar and Secretary.

At the completion of our study, we will make presentations (talks, scientific publications) to other scientists summarizing our findings and their implications. It is hoped our scientific findings will also be disseminated to the general public by coverage in the media.

This project is organised by Professor Lawrence Parsons (Department of Psychology, University of Sheffield), Dr. Renee Timmers (Department of Music, University of Sheffield), Prof. Iain Wilkinson (Department of Radiology, University of Sheffield) Dr. Ian Cross (Department of Music, Cambridge University), and Dr. Tommi Himberg (University of Jyväskylä). It is supported by a grant from the University of Sheffield Arts & Humanities Research Innovation program.

The study was reviewed for ethical participation of volunteers by the Ethics Sub-Committee of the Department of Psychology and Department of Music University of Sheffield.

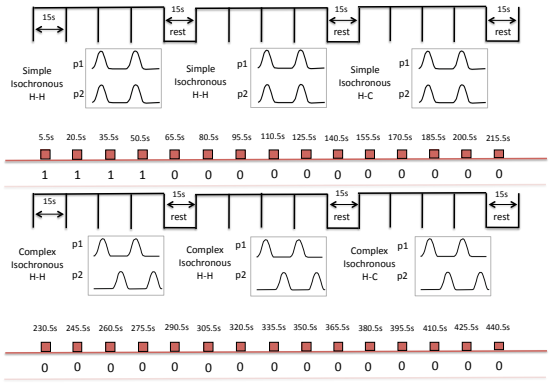
Contact information:
Professor Lawrence Parsons,
Department of Psychology
The University of Sheffield, Western Bank, Sheffield, S10 2TP; 0114 222 6645;
l.parsons@sheffield.ac.uk.

Dr Renee Timmers,
Department of Music, University of Sheffield, 34 Leavygreave Road, Sheffield, S3 7RD; 0114 222 0477; r.timmers@sheffield.ac.uk

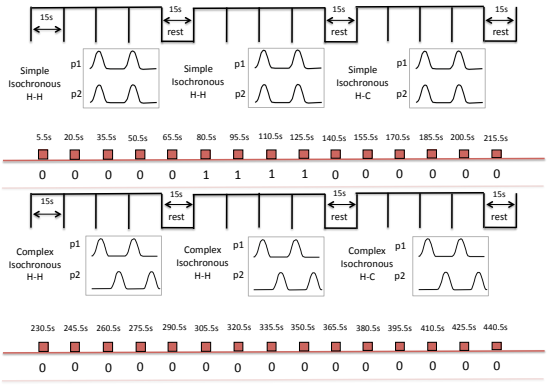
We will be very grateful if you decide to participate. Thank you very much!

APPENDIX B

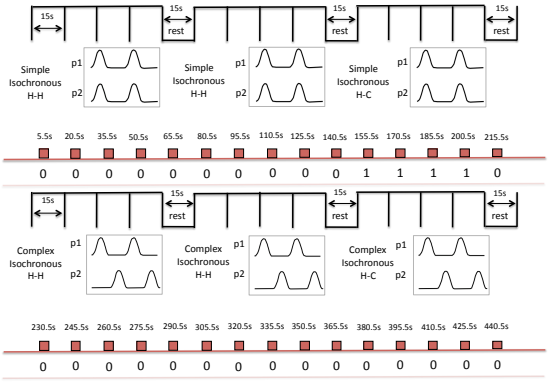
In first level analysis, the six contrasts used for Block 1: Isochronous.



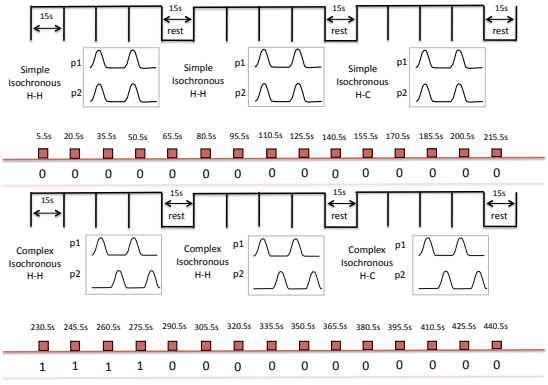
(a) Slide 1



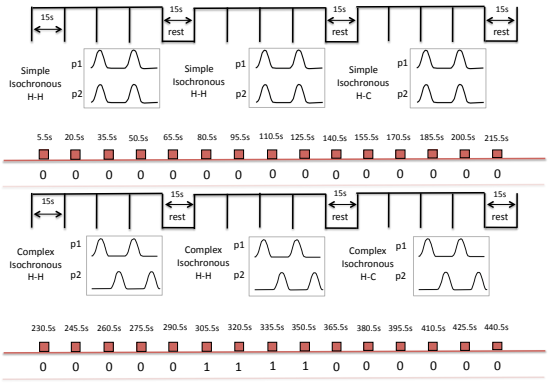
(b) Slide 2



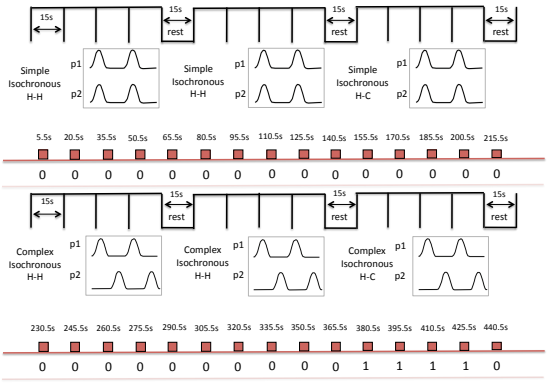
(c) Slide 3



(d) Slide 4

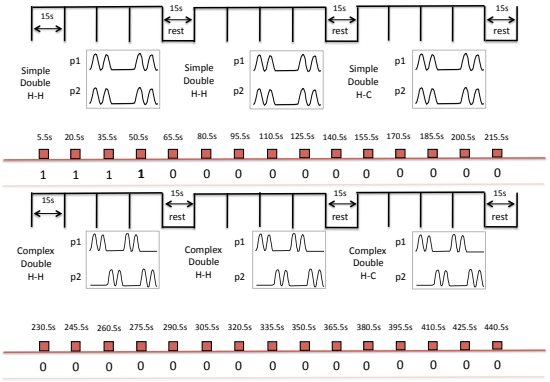


(e) Slide 5

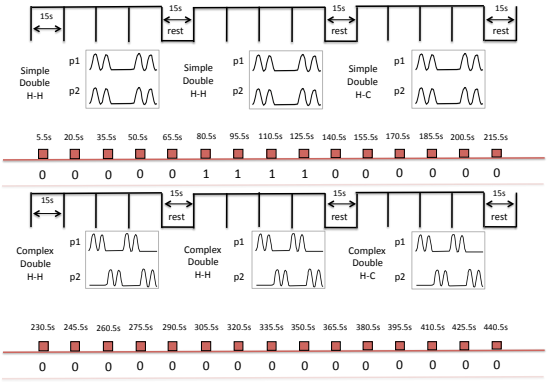


(f) Slide 6

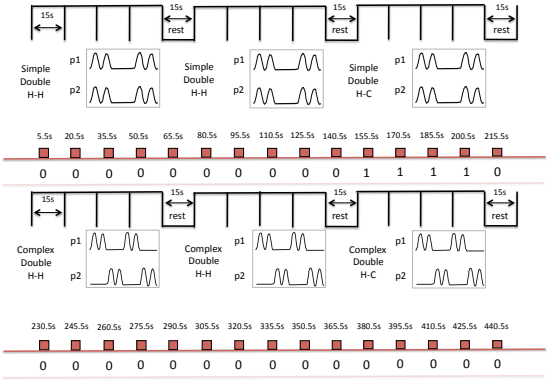
In first level analysis, the six contrasts used for Block 2: Double.



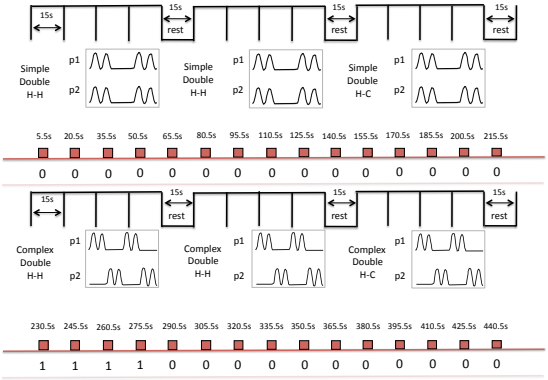
(g) Slide 7



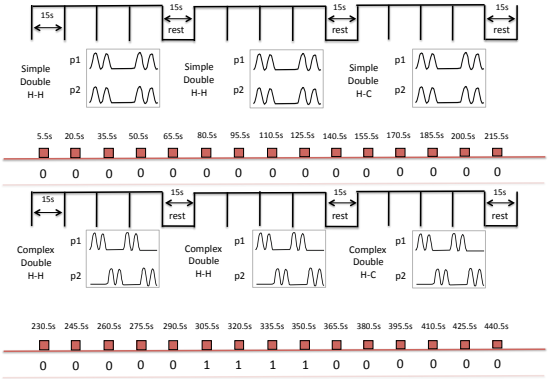
(h) Slide 8



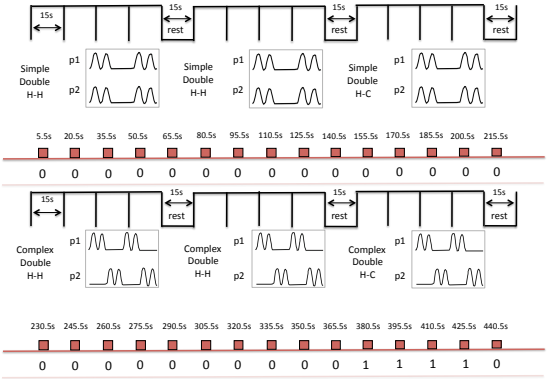
(i) Slide 9



(j) Slide 10

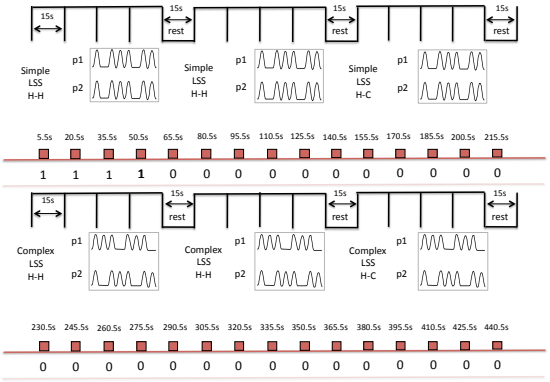


(k) Slide 11

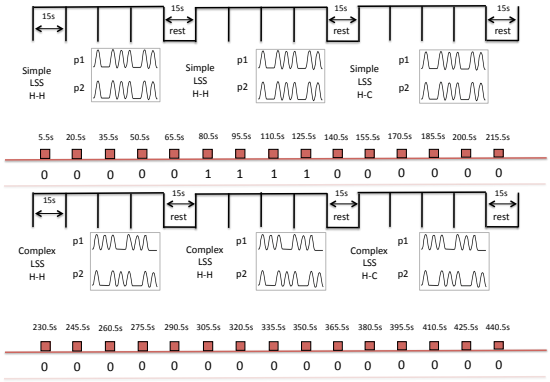


(l) Slide 12

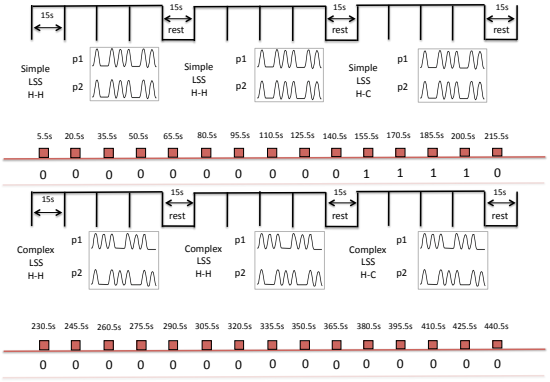
In first level analysis, the six contrasts used for Block 3: Long-Short-Short



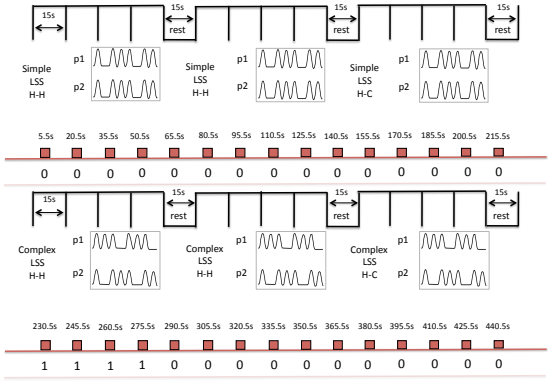
(m) Slide 13



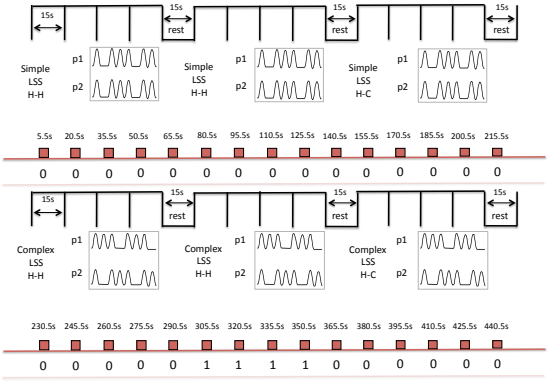
(n) Slide 14



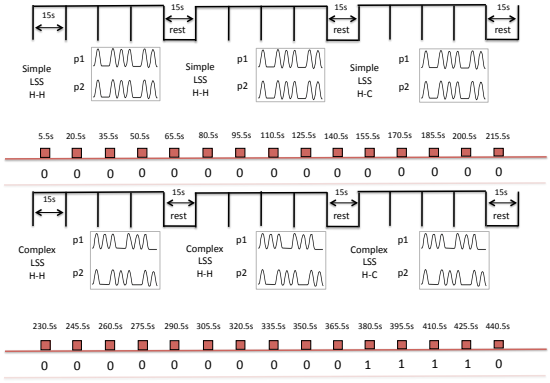
(o) Slide 15



(p) Slide 16

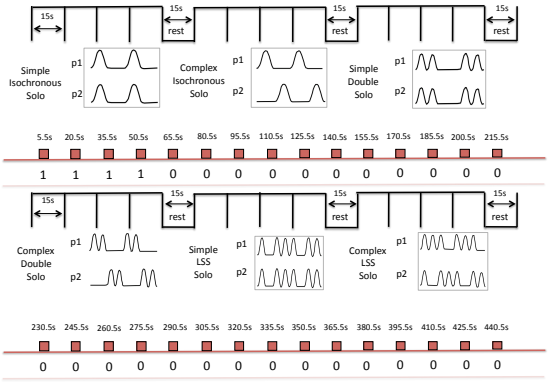


(q) Slide 17

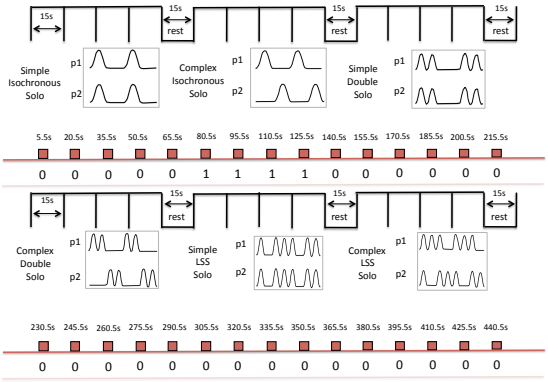


(r) Slide 18

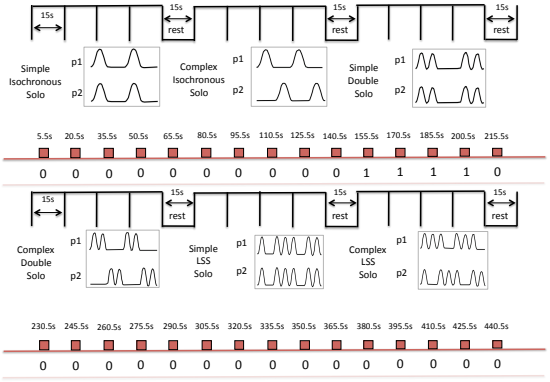
In first level analysis, the six contrasts used for Block 4: Solo



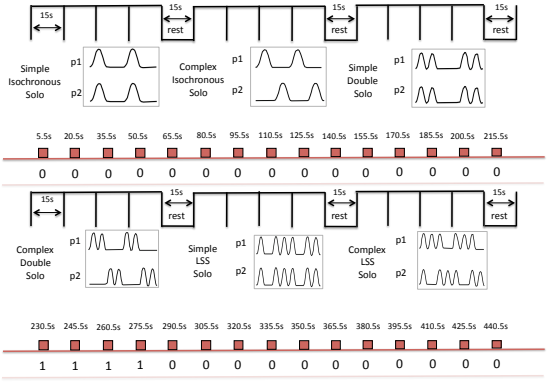
(s) Slide 19



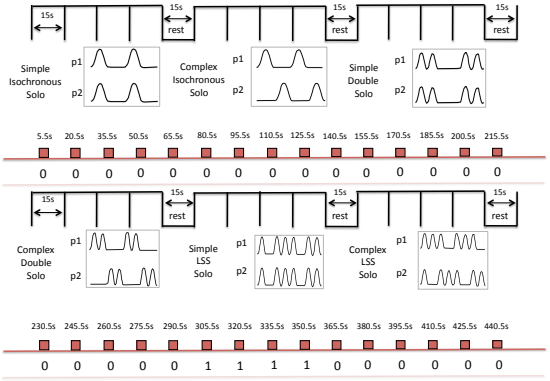
(t) Slide 20



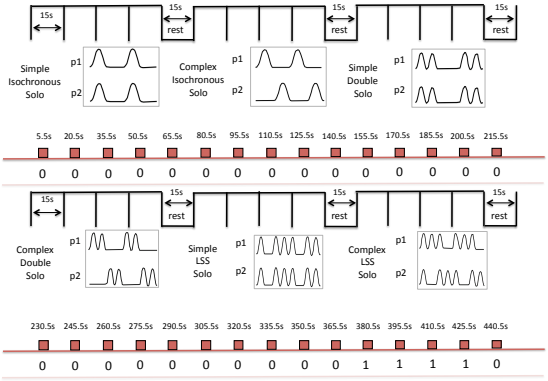
(u) Slide 21



(v) Slide 22



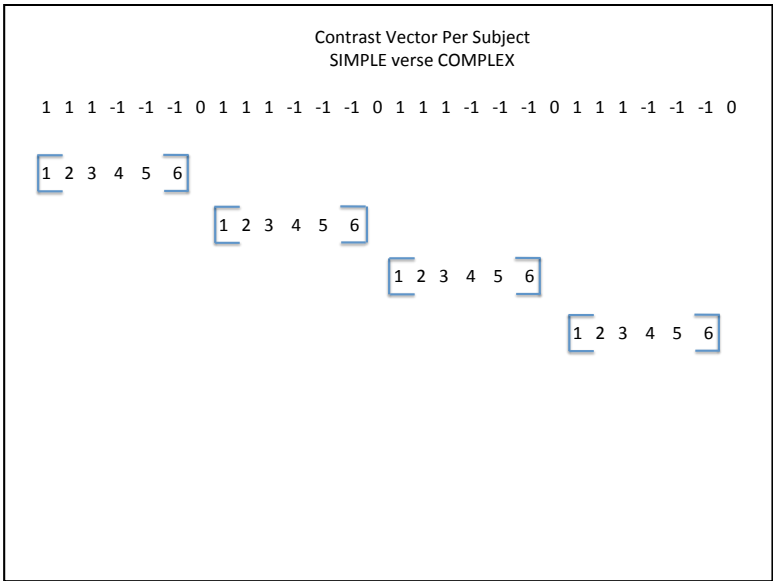
(w) Slide 23

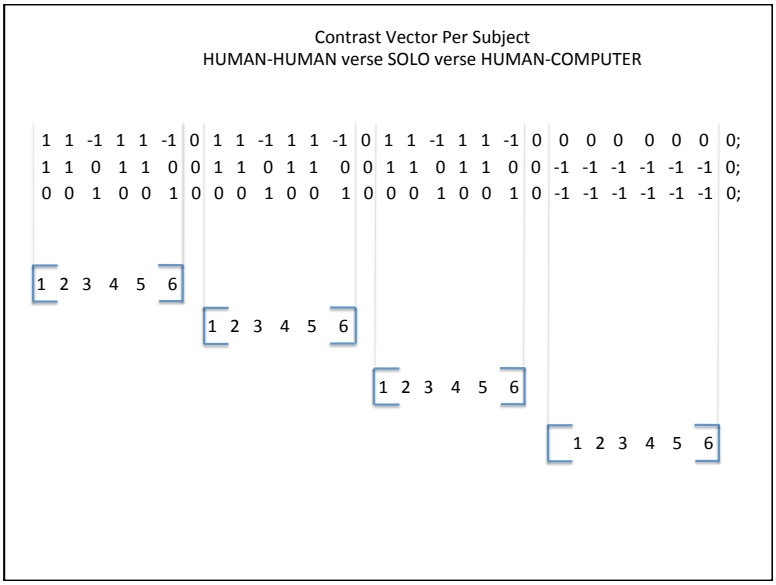


(x) Slide 24

APPENDIX C

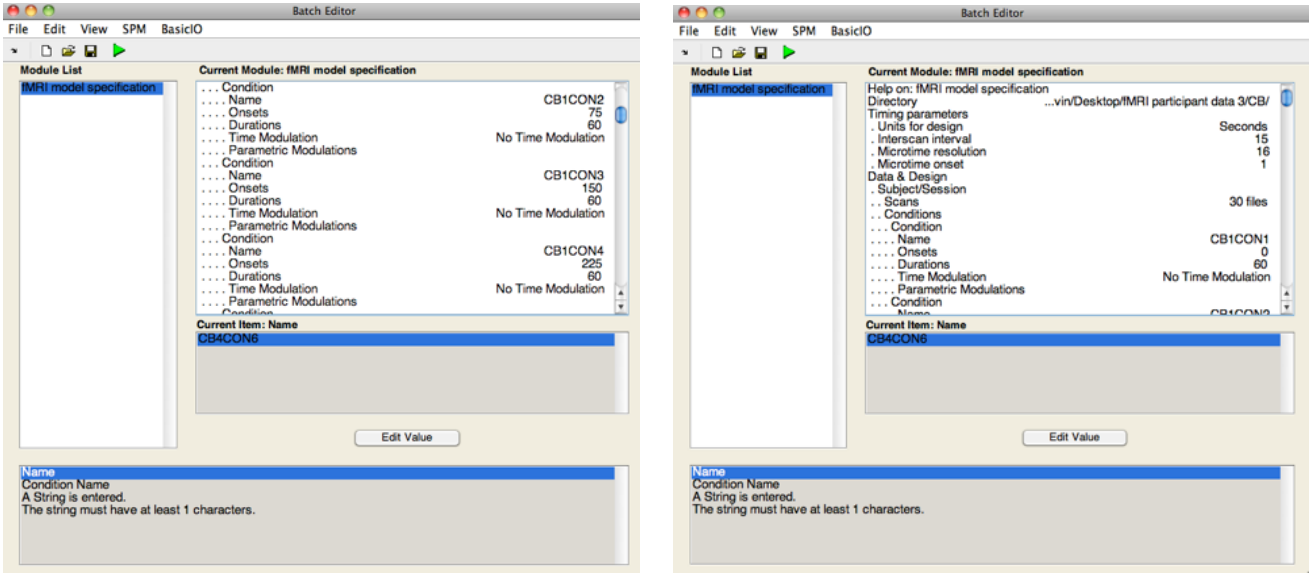
Contrast Vectors



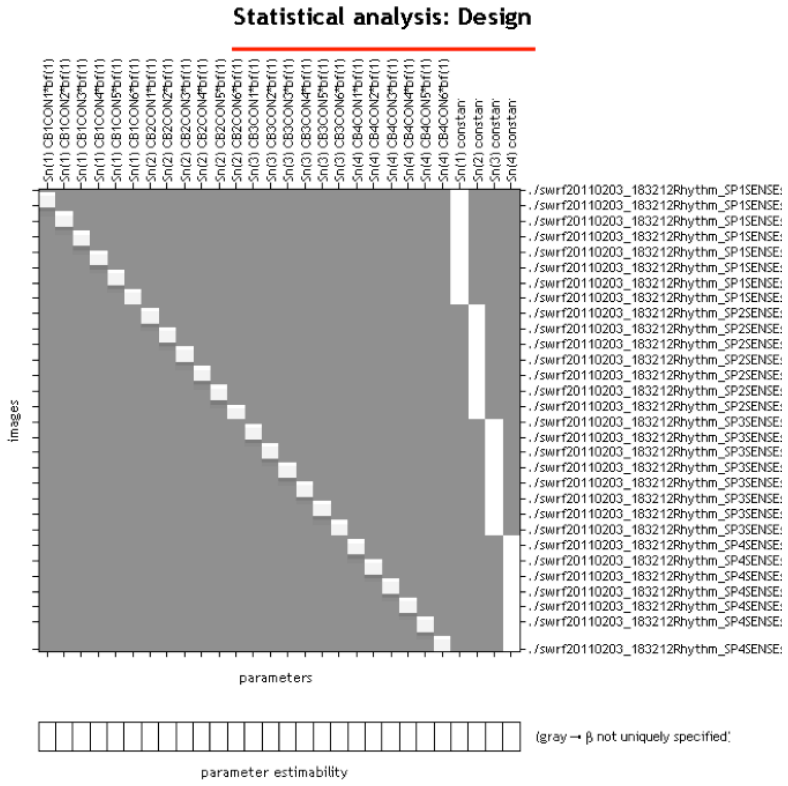


APPENDIX D

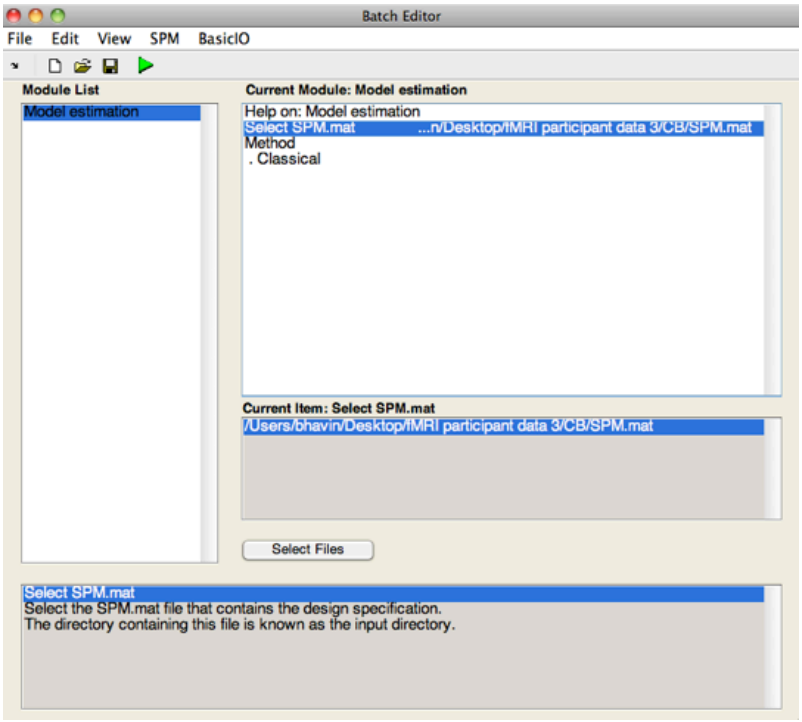
1ST level analysis



Design Matrix

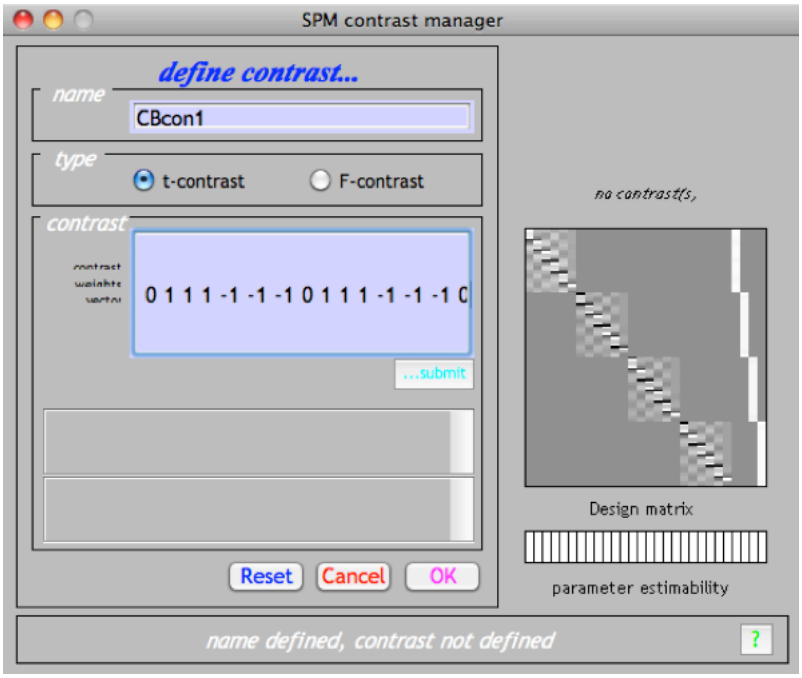


1st Level Estimate

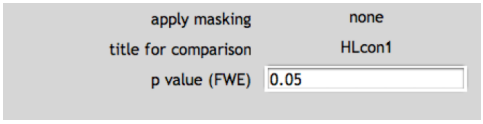


1st Level Results

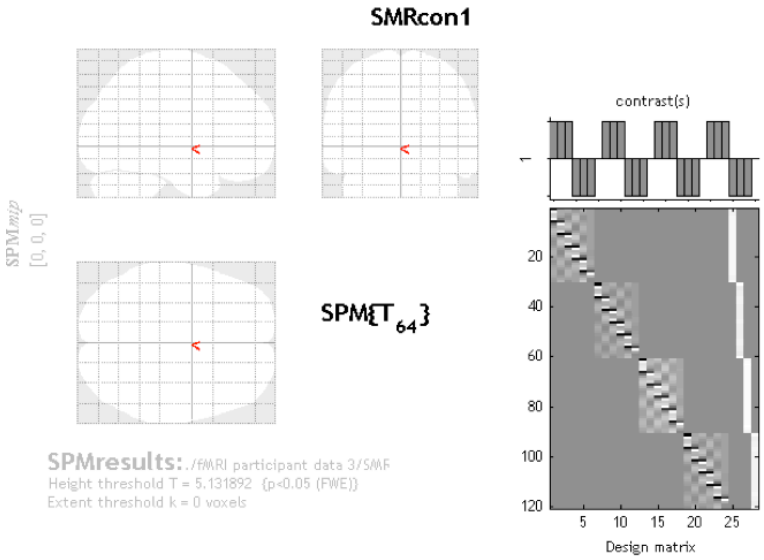
Define contrast for simple verse complex conditions



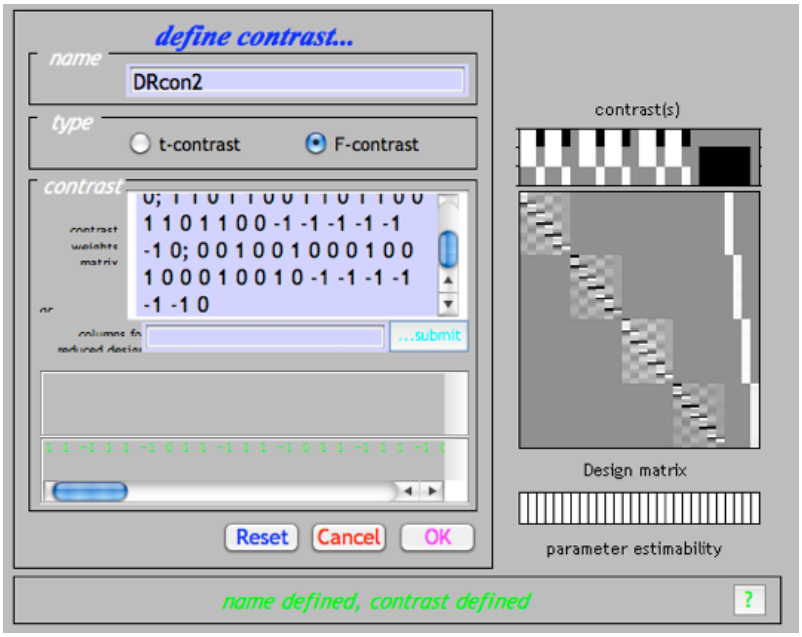
Contrast parameters for simple verse complex conditions



Graphics for simple verse complex conditions



Define contrast for Human-Human verse Solo verse Human-Computer conditions



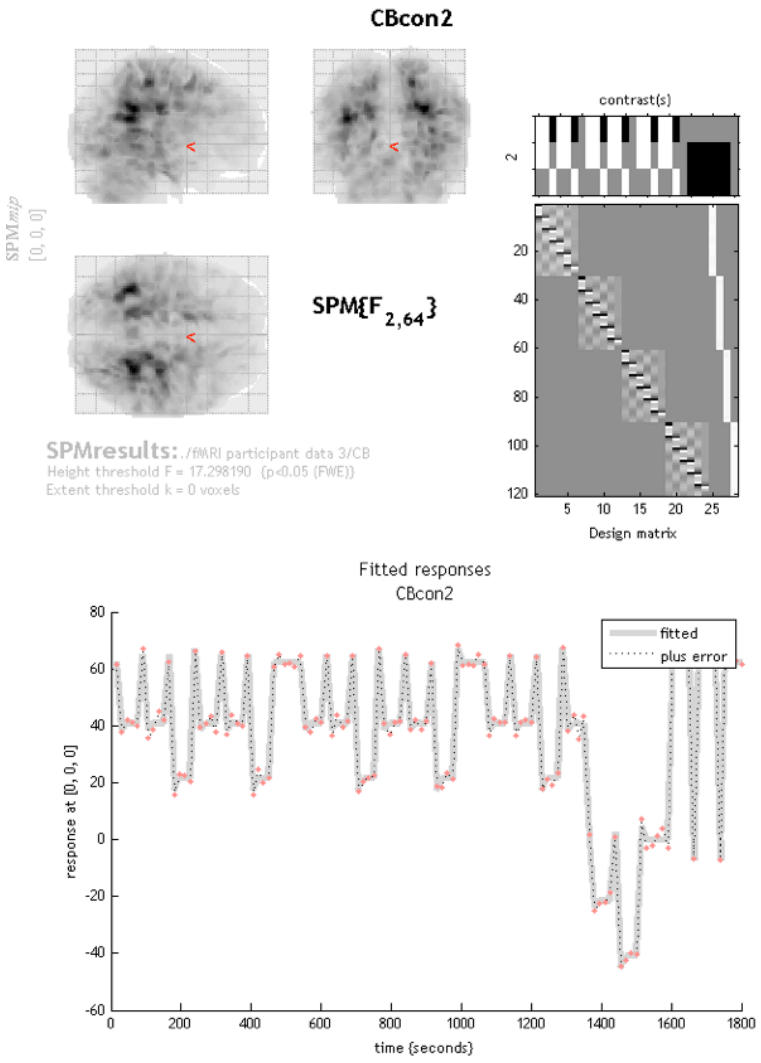
Contrast parameters for Human-Human verse Solo verse Human-Computer conditions

apply masking	none
title for comparison	CBcon2
p value (FWE)	0.05
& extent threshold {voxels}	0

Plotting parameters for Human-Human verse Solo verse Human-Computer conditions

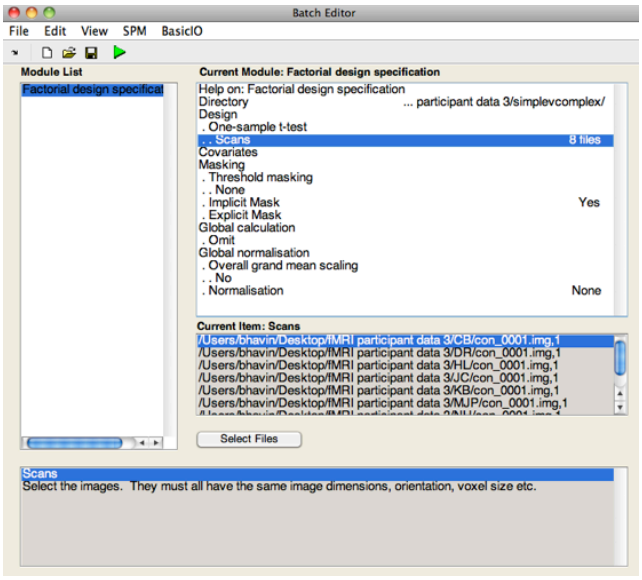
Fitted responses	
CBcon2	
predicted or adjusted response?	predicted
scan or time	

Graphics for Human-Human verse Solo verse Human-Computer conditions

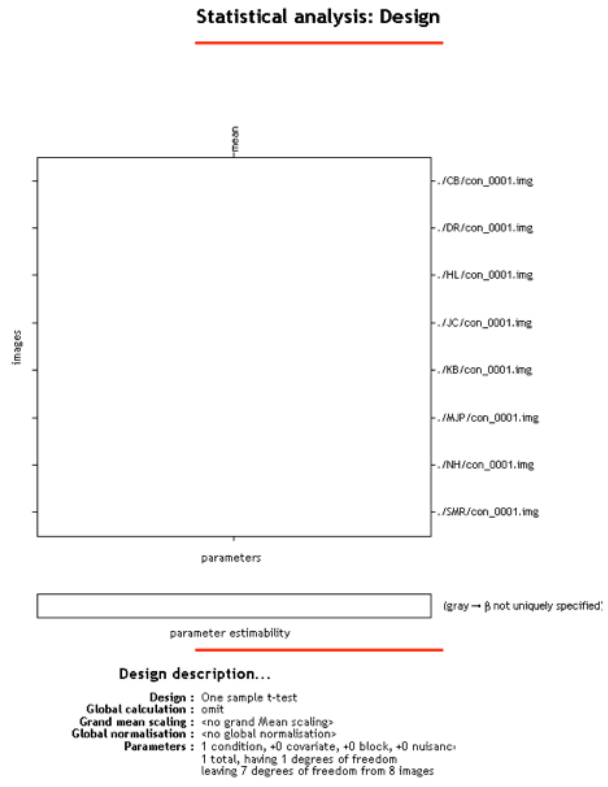


APPENDIX E

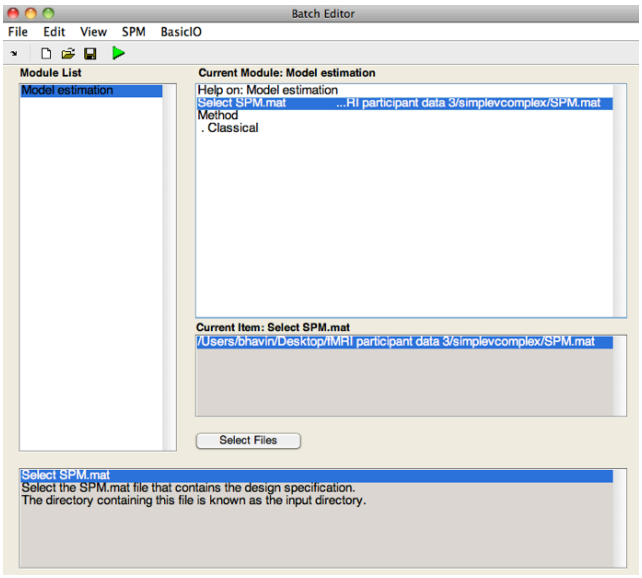
2nd level analysis



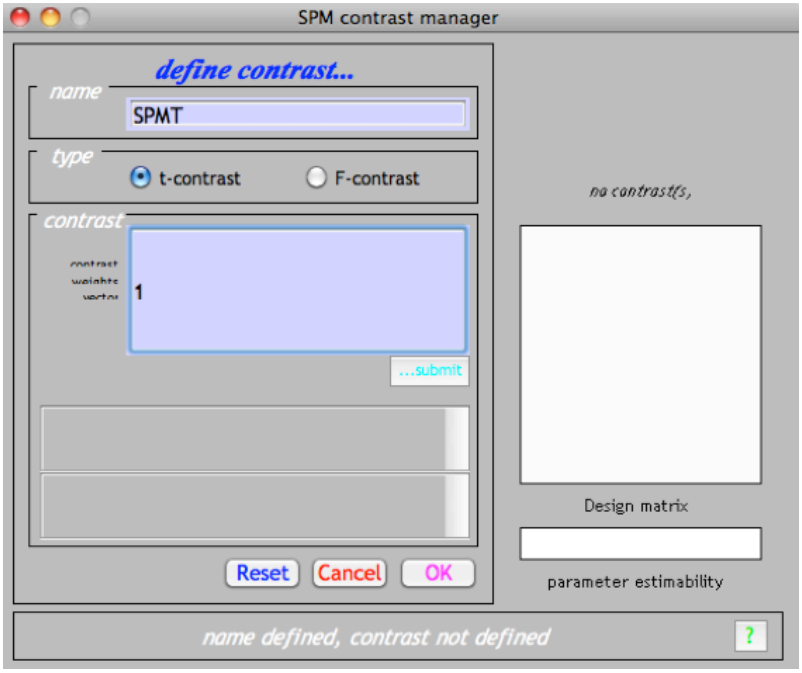
Design Matrix



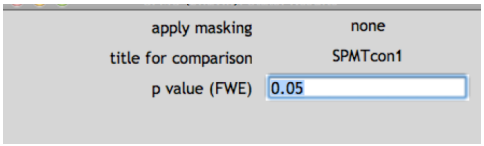
2nd level estimate



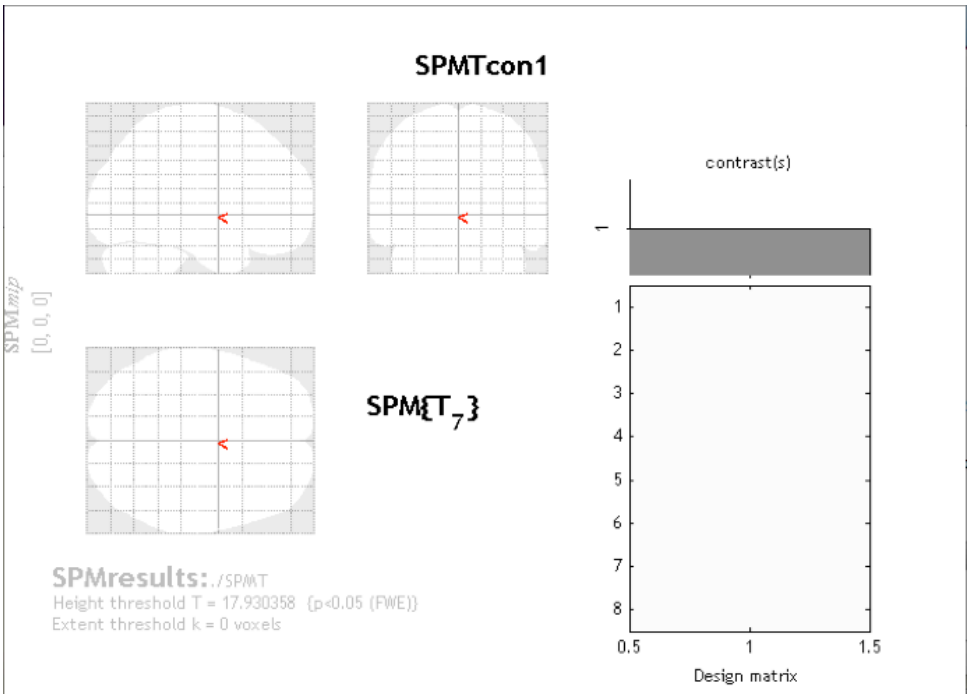
Define contrast for Simple verse Complex condition



Contrast parameters for Simple verse Complex condition



Graphics for Simple verse Complex condition



Statistics: *p-values adjusted for search volume*

set-level		cluster-level				peak-level					mm mm mm
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_\equiv)$	p_{uncorr}	

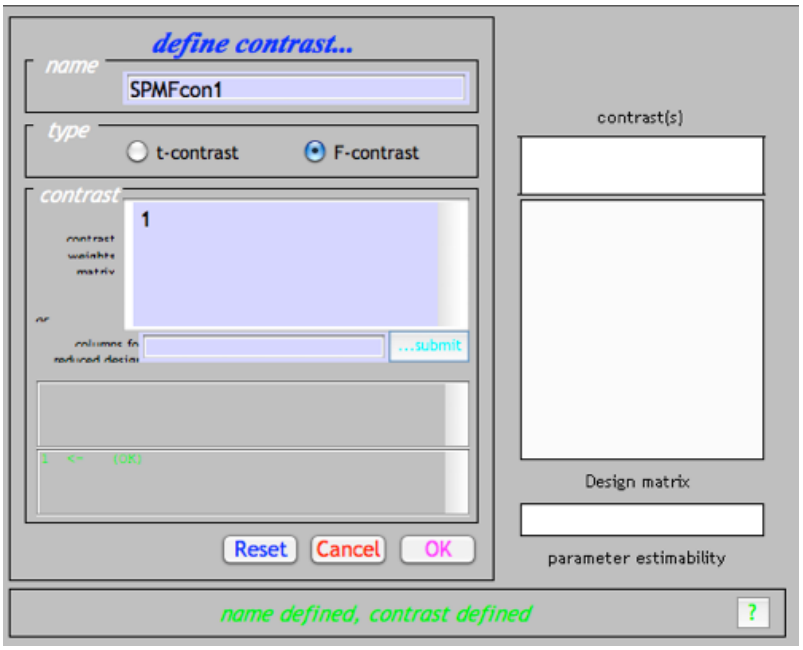
no suprathreshold clusters

table shows 3 local maxima more than 8.0mm apart

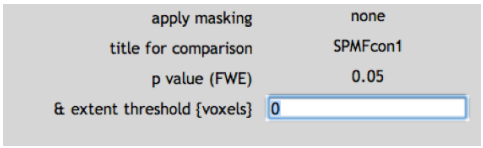
Height threshold: T = 17.93, p = 0.000 (0.050)
Extent threshold: k = 0 voxels, p = 1.000 (0.050
Expected voxels per cluster, <k> = 0.18
Expected number of clusters, <c> = 0.2
FWEp: 17.930, FDRp: Inf, FWEc: Inf, FDRc: Ir

Degrees of freedom = [1.0, 7.0]
FWHM = 13.2 13.2 13.0 mm mm mm; 6.6 6.6 6.5 {voxels}
Volume: 1929952 = 241244 voxels = 795.4 resels
Voxel size: 2.0 2.0 2.0 mm mm mm; (resel = 283.14 voxels)

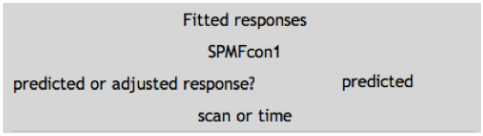
Define contrast for Human-Human verse Solo verse Human-Computer conditions



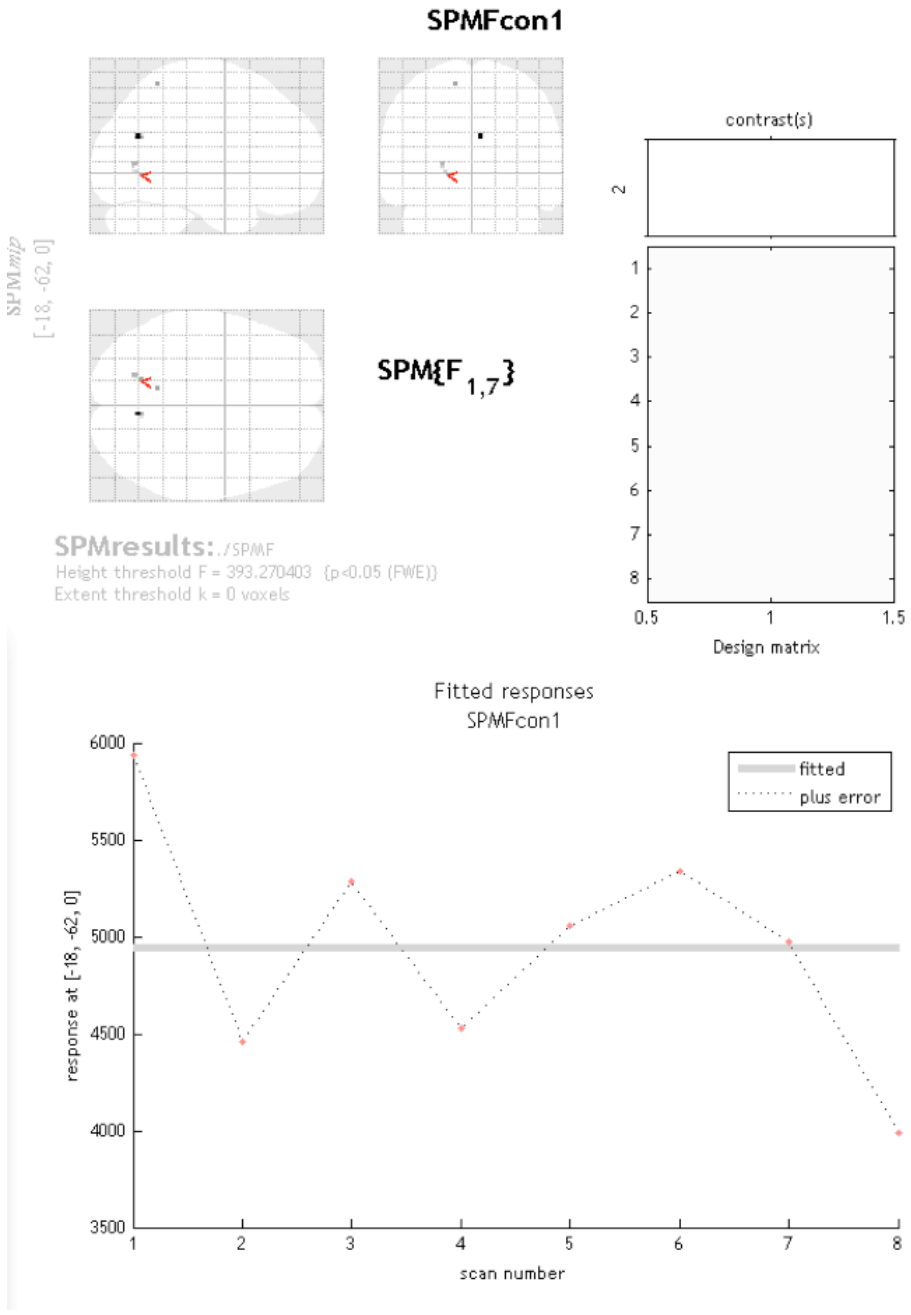
Contrast parameters for Human-Human verse Solo verse Human-Computer conditions



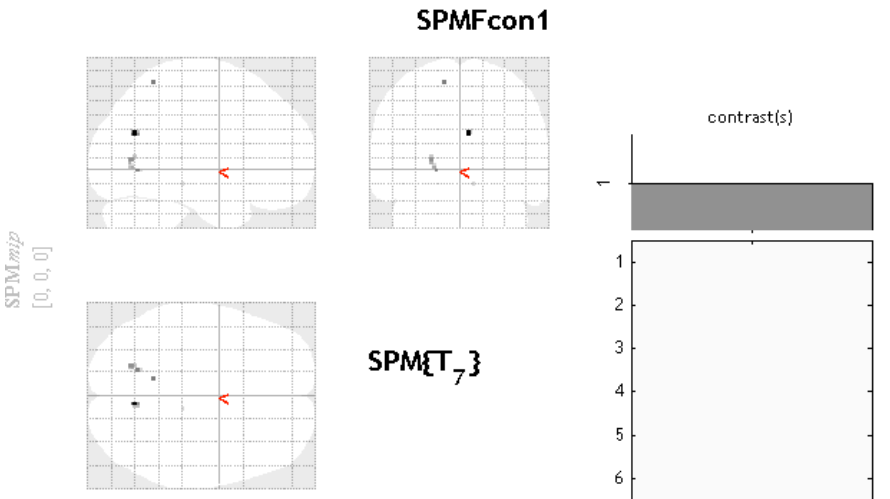
Plotting parameters for Human-Human verse Solo verse Human-Computer conditions



Graphics for Human-Human verse Solo verse Human-Computer conditions



Statistics for Human-Human verse Solo verse Human-Computer conditions



Statistics: ***p-values adjusted for search volume***

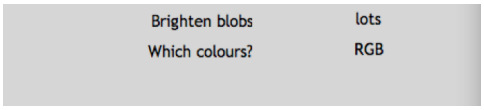
set-level		cluster-level				peak-level					mm mm mm		
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	T	$(Z_\equiv)$	p_{uncorr}			
0.000	5	0.000	0.000	8	0.000	0.002	0.650	28.14	5.63	0.000	6	-64	28
		0.008	0.039	1	0.039	0.009	0.650	23.00	5.38	0.000	-12	-50	64
		0.001	0.010	2	0.006	0.009	0.650	22.93	5.38	0.000	-18	-62	0
		0.000	0.000	8	0.000	0.012	0.650	22.19	5.33	0.000	-22	-66	8
		0.008	0.039	1	0.039	0.038	0.850	18.69	5.12	0.000	10	-28	-10

table shows 3 local maxima more than 8.0mm apart

Height threshold: T = 17.93, p = 0.000 (0.050)
Extent threshold: k = 0 voxels, p = 1.000 (0.050
Expected voxels per cluster, <k> = 0.22
Expected number of clusters, <c> = 0.2
FWEp: 17.930, FDRp: Inf, FWEc: 1, FDRc:

Degrees of freedom = [1.0, 7.0]
FWHM = 14.0 13.9 14.0 mm mm mm; 7.0 7.0 7.0 {voxels}
Volume: 1929952 = 241244 voxels = 663.0 resels
Voxel size: 2.0 2.0 2.0 mm mm mm; (resel = 339.66 voxels)

Parameters for overlay for Human-Human verse Solo verse Human-Computer conditions



Overlap graphics for Human-Human verse Solo verse Human-Computer conditions

