2015

Representation of Affect from FMRI Data as a Function of Stimulus Modality And Judgment Task

Jongwan Kim
University of South Carolina

Representation of Affect from fMRI data as a Function of Stimulus Modality and Judgment Task

by

Jongwan Kim

Bachelor of Arts
Yonsei University, 2006

Master of Arts
Yonsei University, 2008

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Experimental Psychology

College of Arts and Sciences

University of South Carolina

2015

Accepted by:

Douglas H. Wedell, Major Professor

Svetlana V. Shinkareva, Committee Member

Jessica J. Green, Committee Member

Roger D. Newman-Norlund, Committee Member

Lacy Ford, Senior Vice Provost and Dean of Graduate Studies

DEDICATION

To my family.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Douglas Wedell, for his consideration and understanding. He has been my academic father in a new world. I also thank Dr. Svetlana Shinkareva. She has been a guide to explore a new area of research.

I thank the undergraduate assistants in our lab, Rica Karina Buettner, Carly Strohbach, Marian Kate Freeman, Tessa Crawford, and Brenda Nicole Davis, for their help on collecting stimuli, recruiting and running the participants, and help with analysis of behavioral data. I thank the Presidential Fellowship and Laura Griffin Award for the financial support that enabled me to pursue my study and work. I thank my colleague, Wanze Xie, for discussing statistics and having Korean dinners :) I hope my friend, Daehyun Jung, does well in his academic and business career. I also thank my friend, Jeeyea Park, for her support in Korea.

Most importantly, Mom, Dad, and my brother, thank you for your unconditional love and support.

ABSTRACT

The theory of core affect posits that the neural system processes affective aspects of stimuli encountered by the organism quickly and automatically, resulting in a unified affective state described along the dimensions of valence and arousal. Core affect theory posits two functional subsystems that guide affective processing: a sensory integration and a visceramotor network. The proposed study investigates how the representation of affective dimensions depends on sensory modality, features of the task, and brain regions. A series of behavioral studies was run to develop an experimental stimulus set of silent videos and musical clips that met requirements of equating valence across stimulus types while holding arousal constant across valence categories. Valence manipulation was successful, with valence categories were equated on arousal ratings. The stimulus sets in the current study matched many of low level features between valence categories so that any difference between experimental conditions can most likely be attributed to the valence of the stimuli and not to the arousal levels or low level features of the stimuli.

The fMRI study applied multiple multivariate analysis tools to analyze the fMRI data. General valence was successfully decoded from patterns of whole brain activation within participants. The successful cross-modal classification demonstrated that there is modality-general processing of valence at the whole brain level. The multidimensional scaling (MDS) results supported these conclusions by showing that a common valence dimension for visual and auditory trials as well as visual- and auditory-specific valence dimensions. The same analyses were applied to the predefined anatomical ROIs (mPFC,

OFC, and STS) and revealed modality-general valence processing, evidenced by cross-modal classification and the MDS solution. Successful within-participant cross-modal classifications and unsuccessful cross-participant cross-modal classifications implies that modality-general representation of valence could be individual-specific, whereas, successful within-participant within-modal classifications and successful cross-participant within-modal classifications implies that modality-specific representations of valence might be individual-general. A first searchlight analysis was performed to localize the brain regions that were involved in modality-general valence and it identified three significant clusters: right transverse temporal gyrus, left superior temporal gyrus, and right middle temporal gyrus. These searchlight results were validated with cross-modal classification and MDS. The modality-specific regions found by a second searchlight analysis were in the occipital region for visual stimuli and the temporal region for auditory stimuli, as expected. Within-modality classification confirmed that those modality-congruent areas are involved in valence processing of the corresponding modality. Interestingly, each modality's valence was also decoded from the modality-incongruent regions. These results imply modality-specific valence valuation for both modalities in each region, because cross-modal classification was not successful in these regions and MDS did not reveal a general valence dimension in either region.

In sum, the neural representation of both modality-general and modality-specific valence were found at a whole brain level as well as frontal and temporal regions, consistent with the two system approach to core affect posited by Barrett and Bliss-Moreau (2009). This conclusion was bolstered by converging methodologies.

TABLE OF CONTENTS

List of Tables

LIST OF FIGURES

List of Abbreviations

A1 .... Primary Auditory Cortex

ACC .... Accuracy

ANOVA .... Analysis of Variance

EEG .... Electroencephalogram

fMRI .... Functional Magnetic Resonance Imaging

IADS .... International Affective Digitized Sound system

IAPS .... International Affective Picture System

MDS .... Multidimensional Scaling

mPFC .... Medial Prefrontal Cortex

MVPA .... Multivoxel Pattern Analysis

OFC .... Orbitofrontal Cortex

ROI .... Region of Interest

RSA .... Representational Similarity Analysis

RT .... Reaction Time

STS .... Superior Temporal Sulcus

TE .... Echo Time

TR .... Relaxation Time

V1 .... Primary Visual Cortex

CHAPTER 1
INTRODUCTION

The theory of core affect posits that the neural system processes affective aspects of stimuli encountered by the organism quickly and automatically, resulting in a unified affective state described along the dimensions of valence and arousal (Russell, 2003; Russell & Barrett, 1999). Researchers have reported physiological correlates of affective dimensions of experiencing stimuli that differ in valence and arousal levels (Bradley & Lang, 2000; Cacioppo et al., 2000; Kreibig, 2010). Similarly, neuroimaging studies have demonstrated that the affective states corresponding to valence and arousal can be inferred from neural activation patterns (Anders et al., 2004; Baucom et al., 2012; Wilson-Mendenhall et al., 2013). Core affect theory posits two functional subsystems that guide affective processing. The first one is a sensory integration network that affectively encodes sensory information from different modalities and the second one is a visceramotor network that guides autonomic, endocrine and behavioral responses to the objects encountered (Barrett & Bliss-Moreau, 2009). If these two neural systems exist for affect processing, one may ask to what degree affective representations are modality-specific or modality-general for objects and events. One might expect the sensory integration network to process affective information in a modality-specific way but the visceramotor network to process it in a modality-general way. The proposed study investigates how the representation of affective dimensions depends on sensory modality, features of the task, and brain regions.

The traditional approach to the study of affective representation has used univariate statistical methods to determine which brain regions are involved. Due to the multivariate and distributed nature of affective states, pattern-based approaches are well-suited for the study of affect. Pattern-based approaches can evaluate if the whole brain or predefined regions of interest have affect relevant information, can represent the affective states as a lower dimensional space, and allow us to directly compare the similarity representations between regions with theoretical models. Furthermore, they are often powerful enough to apply at the level of the individual rather than rely strictly on group analyses.

This dissertation study used novel multivariate pattern-based approaches and added to our overall understanding of how affective states generated from distinctly different modalities are represented in the brain. I first review the current literature concerning theories of core affect, modality-specificity of affect, and relevant multivariate techniques in Chapter 1. I then introduce methodological approaches of the current study in Chapter 2. Chapter 3 and 4 describe the procedure to develop the stimuli for the main fMRI study. Chapter 3 illustrates the development of stimuli and testing of valence and arousal ratings, and Chapter 4 shows testing of task accuracy and reaction time. In Chapter 5, the main fMRI study is described including methods, procedure, behavioral results, and fMRI results. Chapter 6 summarizes and discusses main findings and suggests future directions.

1.1. Core affect theory

Affect can be represented as a circumplex (Posner, Russell, & Peterson, 2005; Russell, 1980ab), in which affective states are described as falling along a circle in a two-dimensional space. In this model, the most important dimensions are valence and arousal, which constitute the core affective response (Russell, 2003; Russell & Barrett, 1999). Core affect is posited to be a pre-conceptual primitive process, a neurophysiological state, accessible to consciousness as a simple non-reflective feeling (Russell, 2009). It is assumed that the valence and arousal dimensions combine in an integral fashion to form one unified feeling. Arousal refers to a state of being activated or reactive to stimuli, and valence reflects the degree to which the experience is positive or negative. One strength of the core affect concept is that it provides a simple subjective mapping of emotional experience in lower dimensions. The two dimensional representation based on valence and arousal has been successfully used to characterize affective reactions to stimuli in various domains, such as reactions to facial expressions (Abelson & Sermat, 1962; Russell & Bullock, 1985), words (Russel, 1980), and music (Bigand et al., 2005).

Although the two dimensions of core affect provide a powerful tool for representing affective states, it has long been recognized that these two dimensions are not sufficient to fully explain differences in emotional states (Fridja, 1986; Lang, et al., 2008; Roberts &Wedell, 1994). A third dimension often investigated in the framework of core affect is dominance, which can be used to distinguish between negative high arousal states, such as fear and anger. Alternatively, other researchers (Gable & Harmon-Jones, 2008abc; 2010ab) have emphasized the importance of taking into account dimensions related to motivation when considering the consequences of different emotional states. Motivation is typically described by two dimensions: direction and intensity. Motivational direction corresponds to approaching or avoiding a stimulus, which in turn is modified by the dimension of motivational intensity. Although motivational direction is sometimes correlated with valence, researchers operating within this framework consider the two dimensions to be largely distinct. For example, the negatively valenced emotions of anger and fear have opposite motivational directions, with anger reflecting high intensity approach and fear high intensity avoidance. Like fear, disgust and sadness share avoidance motivation, but disgust is at a medium level of intensity and sadness at a low level of intensity. Within the positive valence states, some emotional states, such as comfort, are low intensity approach states and others, such as appetitive states, are high intensity approach states. Harmon-Jones and colleagues (2008abc, 2010ab) have extensively investigated the effect of motivational level on attentional breadth. They found that attentional effects previously linked to valence were more closely related to motivation, so that, for example, disgust and sadness have different effects on attention. In addition to dominance and motivation dimensions, appraisal theory argues there may be additional complex cognitive dimensions that uniquely identify emotions (Roseman et al., 2001; Scherer, 1999; Smith & Ellsworth, 1985). These may include dimensions of control, certainty, risk, etc. Nevertheless, when evaluating emotional responses, the core affect dimensions of valence and arousal typically account for most of the variance (Posner, Russell, & Peterson, 2005; Russell, 1980ab).

1.2. Neural correlates of core affect theory

Neuroimaging studies have supported the core affect dimensional approach by showing that there are separate neural mechanisms for valence and arousal (Anders et al., 2004). Barrett et al. (2007) hypothesized that visceramotor and sensory integration networks may process core affect in brain regions that include the orbitofrontal cortex, anterior insula, amygdala, anterior cingulate cortex, hypothalamus, and ventral striatum. Unlike the belief that amygdala is a ‘fear center’, recent studies have shown that amygdala is more involved in directing the various sources of attention (Holland & Gallagher, 1999) towards a source of sensory stimulation when the predictive value of that stimulation is unknown or uncertain. Thus greater activation of amygdala when participants are exposed to fear-related stimuli may not be a response to fear per se, but a response to potential danger, novelty, or ambiguity (Hsu et al., 2005). Activatioin in the ventral striatum is known to be sensitive to potential gains and losses (Tom et al., 2007). Consistent with this view, both approach and withdrawal behaviors in rats are facilitated via electrical stimulation of the nucleus accumbens, which is part of the ventral striatum. The OFC integrates sensory inputs from the external environment and from the internal body to create a multimodal representation of the environment at a particular moment in time (Mesulam, 2000). It is known that this region is involved in representing reward and threat as well as in hedonic experience (Wager et al., 2008).

There has been an effort to reveal which brain regions are involved in affective processing. Recent meta-analysis studies by Lindquist and colleagues (2012ab, 2015) have concluded that valence generally accompanies increases in activity in bilateral anterior insula, bilateral lateral orbitofrontal cortex (Chikazoe et al., 2014), bilateral amygdala, the ventral striatum, thalamus, dorsomedial prefrontal cortex (BA 9) (Chavez et al., 2014), dorsal ACC, supplementary motor area (BA 6), bilateral ventrolateral prefrontal cortex, and lateral portions of the right temporal/ occipital cortex. Recently, Wilson-Mendenhall, Barrett, and Barsalou (2013) also reported significant correlations between participants' subjective valence and arousal ratings with neural activity in medial orbitofrontal cortex and left amygdala, respectively, both across and within the three emotion categories (happiness, sadness, and fear). They concluded that the neural circuitry utilizes more basic processes across discrete emotions. This result supports the idea that core affect is a basic ingredient of many psychological phenomena, as the affect experienced during discrete emotions shares neural correlates with the affect experienced during simple sensations.

Recent studies have demonstrated lower dimensional representations of affective states based on whole brain fMRI data (Baucom, Wedell, Wang, Blitzer, & Shinkareva, 2012; Shinkareva, Wang, Kim, Facciani, & Wedell, 2014). Multidimensional scaling (MDS) is a set of statistical techniques used to extract underlying dimensionalitywith many MDS studies of affect based on behavioral ratings. Baucom et al. (2012) demonstrated in an fMRI study that the MDS applied to neural responses to International Affective Picture System (IAPS) pictures (Lang, Bradley, & Cuthbert, 1999) produced the same affective circumplex found for behavioral judgments of similarity. Valence was also decoded by the same research group (Shinkareva et al., 2014) using IAPS picture and IADS (International Affective Digitized Sounds) (Bradley & Lang, 1999) sound stimuli. In both studies, active voxels that were predictive of affective states in the MDS analysis and multivoxel pattern analysis were distributed across the whole brain and analyses by region of interest (ROI) did not implicate any particular region, supporting the idea that processing of core affect may not be localized but distributed. All in all, these combined results imply that the core affect dimensions of valence and arousal provide a stable basis for characterizing the processing of affective stimuli. The current study will further investigate if these relationships are tied to specific ROIs using various methodologies.

1.3. Modality-specific vs. modality-general processing of affect

Barrett and Bliss-Moreau (2009) have posited that core affect may be supported by two neural systems, a sensory-integration network and a viscera-motor network. A sensory-integration network establishes an experience-dependent, value-based representation of an object that includes both sensory features of an object and its effect on the homeostatic state of the body, while a viscera-motor network is part of a functional circuit that guides autonomic, endocrine, and behavioral responses to an object. The brain regions involved in viscera-motor network modulate changes in the viscera associated with the autonomic nervous system. It follows that affective coding in the sensory-integration network may be modality-specific, while processing in the viscera-motor network might be modality-general in nature. Modality-specific valence encoding implies that neural responses to valence of stimuli from one sensory modality do not correspond to the neural responses to valence from another modality. Modality-general valence encoding, on the other hand, implies that the same neural responses occur to valence values regardless of the modality of the eliciting stimuli.

Recent research has examined the neural representations of affect elicited by different stimulus modalities. Shinkareva and colleagues (2014) manipulated the valence of pictures and sounds presented to participants in an fMRI study designed to determine if encoding of affective states is modality-specific or modality-general. Using multivariate pattern analysis (MVPA) classification methods applied to individual participants, they found that valence for each modality could be predicted from a classifier trained on the same modality trials, within-modality classification, but not from a classifier trained on other modality trials, cross-modality classification. Moreover, significance tests based on whole brain contrasts found evidence for modality-specific valence encoding but not for modality-general valence encoding in each of the participants. Application of multidimensional scaling to similarity of activation patterns for the different conditions resulted in a solution that was again consistent with modality-specific processing. The multidimensional scaling solution revealed separate valence dimensions for each modality, but a common valence dimension for both modalities was not extracted. These results suggest modality-specific processing of affect.

However, other neuroimaging studies have yielded evidence for modality-general processing. Peelen, Atkinson, and Vuilleumier (2010) found evidence of modality-general processing across five emotion conditions (anger, disgust, fear, happiness, and sadness) portrayed in either videos of facial expression, videos of body expressions, or auditory recordings of vocalized nonlinguistic expressions. They found evidence of common emotional processing located in the medial prefrontal cortex and the left superior temporal sulcus.

Studies have also found significant cross-modal adaptation aftereffects between face and voice stimuli, which supports the modality-general processing of affect (Pye & Bestelmeyer, 2015; Watson et al., 2014). In particular, Watson et al. (2014) found within-modal adaption effects in sensory cortices whereas cross-modal adaption effects were found in the posterior superior temporal sulcus, suggesting that the STS is involved in multimodal integration regardless of the affective content of the information. Taken together with the previously cited results, the determination of modality-specific and modality-general processing should be highly tied to ROI analyses.

What causes the difference between modality-specific and modality-general processing of affect? One possibility is that modality-general affective processing may be engaged whenever a response related to affect or emotion must be generated by the participant. The Shinkareva et al. (2014) study that found only modality-specific processing used an incidental exposure paradigm with no judgment task. The Peelen et al. (2010) study that found modality-general processing used an explicit emotional judgment task that should have prompted engagement of the more general viscera-motor network. A second difference between the studies concerns the stimuli themselves. Shinkareva et al. (2014) manipulated valence whereas Peelen et al. (2010) manipulated more discrete emotions. It may be that a more general response is elicited for discreet emotions. A recent study by Chikazoe, Lee, Kriegeskorte, and Anderson (2014) examined two stimulus modalities (visual pictures and gustatory tastes) that varied in valence and were rated by participants in an fMRI scanner. They found evidence for both modality-general processing in the orbitofrontal cortex and modality-specific processing in the ventral temporal cortex (visual) and anterior insular cortex (gustatory). This study shows that modality-specific processing in some regions of the brain occurs even with explicit judgment of affect. Furthermore, it indicates that modality-general processing of affect can occur with manipulations of valence and not just with manipulations of specific discrete emotions. The proposed research builds on these findings and directly investigates the role of making overt affective-based responses to stimuli versus responses that are not related to affect.

One interesting research question is the relationship between sensory region and its ability to process affective information. Though frontal areas like OFC or mPFC (Chikazoe et al., 2014; Etkin et al., 2011; Said et al., 2010) and subcortical areas like insula or amygdala are known to be involved in affective processing, it has been reported that sensory areas are also involved in modality-congruent affect processing. For example, patterns of activity in voice-sensitive cortices can be used to distinguish categorical emotional vocal expressions (Ethofer et al., 2009). Similarly, Lang and colleagues (1998) presented emotion-inducing pictures and found a greater functional activity for emotional pictures than for neutral pictures in primary visual regions, including occipital gyrus, fusiform gyrus, and lingual gyrus. Thus, successful classification of valence is expected in the modality-congruent regions. However, there is reason to believe that modality-incongruent regions may also have valence information. Meyer et al. (2010) demonstrated successful decoding of silent visual stimuli from primary auditory cortex. Furthermore, when participants imagined visual objects in the complete absence of perceptual input, primary visual cortices were activated and appeared to specifically represent the contents of the participants' visual experience (Kosslyn et al., 1995). These findings suggest that sensory areas may encode perceptual experience rather than perceptual input *per se*, and decoding affective stimuli could be successful in a modality-incongruent regions. This will be tested by training and testing classifiers from visual trials in auditory cortex, and vice versa.

1.4. Modulation of attentional focus on affective response

It is important to consider the effect of attentional focus on affective responses, oftentimes manipulated by comparing explicit and implicit tasks (Cunningham et al., 2004; Hutcherson et al., 2005; Lange et al., 2003; Straube et al., 2004). Pessoa et al. (2002) demonstrated that exposure to affective stimuli without attentional resources resulted in less activation of amygdala, suggesting that the processing of facial expression appears to be under top-down control. Typically, these studies compared the brain activation levels when participants were exposed to affective stimuli in a passive viewing paradigm and when they were asked to rate affective responses on scales during or after stimulus presentation. For example, Hutcherson and colleagues (2005) compared neural responses to amusing and sad videos with continuous self-report ratings or passive viewing conditions. They found that rating condition produced increased activity in anterior cingulate, insula, and several other areas associated with introspection of affect. Lange and colleagues (2003) compared an explicit affect rating task to an irrelevant task (gender decision of figures in stimulus) and to no task (passive viewing). They found greater neural activity in ventral prefrontal cortex in the affect rating condition compared to the other task conditions. Similarly other researchers have found a greater activation in affect-relevant brain regions for explicit tasks compared to implicit tasks (Cunningham et al., 2004; Straube et al., 2004). In sum, previous studies have suggested that compared to implicit tasks, focusing on the affective aspect of stimuli may increase the affective responses in the brain.

Given that one of the two-neural systems (Barrett & Bliss-Moreau, 2009) of affective processing is sensory dependent and the other is not, the modality-specificity of affective processing may depend on the attentional focus. Studies that reported modality-general processing mentioned in the previous section (i.e. Chikazoe, et al., 2014; Peelen et al., 2010) asked participants to rate their affective responses after stimulus presentation, so it is possible that rating affective response may have led them to focus more on the affective aspect of stimuli regardless of stimulus modality. For example, Mothes-Lashe and colleagues (2015) presented threatening and neutral prosodies with a perceptual task and participants were asked to determine either the gender of the speaker (auditory task) or the kind of visual symbol (visual task). They found a significant difference of amygdala activations between threatening and neutral prosodies only during the auditory task but not during the visual task, suggesting that activation of amygdala to threat-related voices depends on modality-specific attention. In sum, it can be hypothesized that the nature of affective processing may depend on the attentional focus (general affective aspect versus non-affective aspect). Specifically, focusing on the affective aspect of the stimuli may lead to modality-general processing, whereas focusing on the non-affective (perceptual or semantic) aspect may lead to modality-specific processing of affect.

CHAPTER 2

APPROACHES TO THE STUDY OF REPRESENTATION OF AFFECT

2.1. Analytical approaches

2.1.1. Multivoxel pattern analysis

The traditional statistical method for analyzing neuroimaging data is univariate based. For example, statistical parametric mapping treats individual voxels independently with separate general linear models. One disadvantage of the univariate approach is the issue of multiple comparison. Familywise error correction can be used but it greatly reduces the statistical power. One way to reduce this problem is to focus on specific regions of interests. For example, one may examine voxels only from the amygdala or striatum for emotion studies so that a smaller number of voxels are evaluated compared to the whole brain. As a result, familywise correction will be less severe. However, an ROI based approach may not be used when one lacks theoretical and empirical knowledge of relevant ROIs. If relevant processing is postulated to be in distributed networks across the whole brain, the ROI approach may also fail. For example, Baucom et al. (2012) successfully decoded affective states of valence and arousal using distributed voxels over the whole brain rather than voxels from specific ROIs. Finally, the univariate approach may not be able to capture the joint activity patterns of multiple voxels in the whole brain or in ROIs, like when different neurons within the same brain region perform different tasks. For example, the mean activation within an ROI might be the same for two experimental conditions, but the patterns of activation within that ROI for each condition might be different.

In these cases multivariate analyses can increase statistical power. Another advantage of multivariate analyses is related to the nature of affect. Though it is known that affective information may be localized in multiple brain regions like the amygdala or insula, it has been also argued that valence may be represented in a more distributed pattern (Baucom et al., 2012) or by a flexible set of valence-related regions (Lindquist et al., 2015). As discussed above, there is convincing evidence to show that affective states can be mapped on the lower dimensional spaces of valence and arousal. Thus some statistical techniques, like multidimensional scaling (MDS), might be appropriate for revealing the nature of affect.

Multivoxel pattern analysis (MVPA) has been proposed to take into account the full spatial pattern of brain activity in fMRI studies (Cox & Savoy, 2003; Haxby et al., 2001; Haynes & Rees, 2006). Even though some cognitive states are encoded in spatially distinct locations in the brain, such as the fusiform face area for face processing and the parahippocampal place area for houses and visual scenes, a wide variety of perceptual or cognitive tasks can be decoded using analyses of spatially distributed rather than localized patterns of brain activity.

MVPA techniques have been also used in other domains including EEG for face (Cauchoix et al., 2014), sound (Brandmeyer et al., 2013), music (Schaefer et al., 2011) and eye movements for cognitive tasks (Henderson, Shinkareva, Wang, Luke, & Olejarczyk, 2013). For example, Henderson and colleagues (2013) used the multivariate pattern classification technique to predict four types of cognitive tasks (scene search, scene memorization, reading, and pseudo-reading) from four measurements of eye movements (means and standard deviations of fixation duration and saccade amplitude). The classification was performed in a leave-one-out cross validation approach within the same experimental session and across sessions. The result showed that classification accuracies within and across sessions were significantly higher than the chance level (.25), and the two tasks (scene search and scene memorization) that involve the identical stimuli were also well discriminated from each other (above chance). In sum, Henderson et al. (2013) demonstrated robust classification accuracy both within- session and across-sessions. Please note that this study only used four features (M and SD of fixation duration and saccade amplitude), which is different from EEG and fMRI MVPA studies. In fMRI, each voxel response serves as ‘one’ measurement so that there are thousands of measurements in fMRI MVPA studies. The Henderson et al. (2013) study demonstrated a successful classification from only a few measurements.

The pattern classification approach to physiological response to emotion can be complementary to traditional group mean approaches. In pattern classification analysis, units consisting of multiple measures are classified into predetermined affective states. A participant’s affective state is predicted from a combination of physiological features (Kreibig, Brosch, & Schaefer, 2010). Previous studies heavily relied on the linear classifier approach such as linear discriminant analysis. For example, Kreibig and colleagues (2007) measured galvanic skin response, heart rate, and respiration parameters while participants were viewing fearful or sad film clips and performed linear discriminant analysis. They found substantial accuracy (78.6% - 89.3 %) depending on the subsets of physiological parameters. Some studies, but not all, included cross validation techniques (leave-one-out). Kolodyazhniy and colleagues (2011) argued that nonlinear models were systematically better than the linear one in all four cross-validation settings, with only five common features out of fourteen, suggesting that even though there is a slight difference between linear and nonlinear models in terms of classification accuracy, the nonlinear models were able to provide the similar accuracy with a smaller number of physiological features.

Different types of stimuli have been used for multivariate classification studies involving affect. Coutinho and Cangelosi (2011) used musical stimuli to induce affective states. This study successfully demonstrated the use of both low-level properties and physiological variables to predict the affective state of participants, showing the possibility that multiple sources of features can be used simultaneously in other domains (i.e. voxel responses or eye-movement data) using various modality features (visual– hue, saturation, brightness, visual complexity or motion– movement speed). Russo, Vempala, and Sandstrom (2013) used musical stimuli and tested both linear and nonlinear approaches for statistical analysis. The linear modeling approach (multiple regression) significantly predicted only arousal, not valence, but the nonlinear approach (neural network modeling) successfully predicted both valence and arousal. In sum, these studies suggest that the low level features of the stimuli should be taken into account in the neuroimaging study.

2.1.2. Similarity-based analysis

Similarity based approaches provide another way to understand fine-grained voxel pattern information. These approaches include the examination of the similarity structure between items using multidimensional scaling (MDS) (Shinkareva, Wang, & Wedell, 2013). MDS and related techniques can also be used to examine the 'similarity between similarity structures' of different brain areas (i.e. voxel responses of primary visual cortex vs. auditory cortex), species (men vs. monkeys), or different types of responses (brain activity measurement, behavioral measurement, psychophysiological measurement, and computational modeling) (Kriegeskorte, Mur, & Bandettini, 2008).

The basic assumption of MDS on fMRI data is that a series of voxels will show a similar pattern of activation for same-category items (i.e. the two exemplars for the positive condition, or the two exemplars for negative condition) and will show a dissimilar pattern of activation for different-category items (i.e. a positive exemplar and a negative exemplar).

There are two primary ways to collect the similarity data in behavioral studies: direct and indirect. The direct method asks participant to rate the similarity of each pair of stimuli. For example, in a face perception study, two face pictures are presented simultaneously and participants are asked to rate the perceived similarity between the two pictures. Suppose a pair of 'sad' and 'disgust' is rated high and a pair of 'sad' and 'happy' and another pair of 'disgust' and 'happy' are rated low. Then 'sad' and 'disgust' are rated as similar each other whereas these two emotions are dissimilar to 'happy'. Low level representation of these three emotions using multidimensional scaling will show that the first two emotions will be placed closely together, whereas the last emotion will be placed far from the first two.

One of the disadvantages of the direct rating is the number of trials. As the number of items increases, the number of trials increases exponentially. Let k be the number of the items, then the number of trials is $k*(k-1)/2$. The second disadvantage is that direct similarity ratings can be collected from only human participants behaviorally, which means that similarity ratings cannot be collected in animal or other human studies (psychophysiology/neuroimaging) in which pairwise assessment is not made. There are many indirect ways to infer similarity based on the profiles' response, confusion data, etc. Based on multiple measures, data profiles are constructed for each item. The data profile can be behavioral ratings, psychophysiology measures, or neural activations. Then the profile-by-item data matrix is correlated, after which then the resulting correlation coefficients serve as a proximity measure. This correlation matrix can be subtracted from 1, and then the resulting matrix is a dissimilarity matrix ranging between 0 and 2 (alternatively, distance measures can be computed directly from the profiles). The distance matrix can be directly compared to other types of dissimilarity or distance matrices.

In the neuroimaging studies of affective states, Shinkareva and colleagues have used similarity based analyses like INDSCAL (Baucom et al., 2012) and STATIS (Shinkareva et al., 2014) to investigate the affective space from the functional pattern of whole-brain activity elicited by viewing pictures or listening to sounds. Baucom et al. (2012) successfully extracted a two-dimensional space of valence and arousal using pictures, which was consistent with the representation from behavioral responses. Shinkareva et al. (2014) derived a three dimensional representation that coded modality, picture valence and sound valence, consistent with modality-specific processing. An EEG data set also illustrated a valence by arousal representation (Onton & Makeig, 2009). Finally, Kim and colleagues (2015) explored the brain regions that have associated emotion representation and demonstrated clustered emotion items in posterior cingulate cortex, mPFC, precuneus, and angular gyrus using fMRI data from viewing video clips.

2.2. The current study

The purpose of the current study is to explore the modality-general or modality-specific representation of affective states under two types of the tasks. Analytical approaches described above will be used for the current study.

2.2.1. Within-participant whole-brain MVPA

To determine if patterns of whole-brain activity elicited by exposure to affective stimuli can be used to predict the affective states, MVPA will be performed on the whole-brain patterns of activity for each participant following procedures used in prior work (Baucom et al., 2012; Shinkareva et al., 2014). The data were divided into a test set, containing one run, and a training set, containing the other runs of trials.

To reduce data size, feature selection will be performed on the training data, choosing the most stable voxels across multiple presentations of a condition. Voxel stability scores are computed by averaging pairwise correlation coefficients between vectors of presentations of all conditions in the training set, thus assigning higher scores to voxels with more stable variation in activity across conditions in the training set. Voxels are then ordered within each individual's data set from highest to lowest stability. As past research suggests greatest stability in the 400-600 voxel range (Baucom et al., 2012), 600 voxels will be investigated for whole brain analysis[1].

[1] Within-participant general-valence 3-way classifications (positive vs. negative vs. neutral) were performed based on 9 numbers of voxels (50, 100, 250, 400, 600, 1000, 2000, 3000, and 4000). The classification accuracies increased from the lowest number of voxels (50) to 400 voxels and stayed consistent through the largest number of voxels (4000). Another test was conducted by 2-way classifications (positive vs. negative,

A logistic regression classifier will be trained using a leave-one-fold-out cross-validation technique. In cross-modal classification, one fold contains only visual trials and the test fold contains only auditory trials, or vice versa. This cross-modal classification should only be successful if the voxels contain modality-general valence information. In within-modality classification, only visual trials or only auditory trials are trained and tested separately. This approach provides the most powerful test for valence information in each modality, but it does not distinguish between modality-specific and modality-general representation. Modality can also be ignored in classification, testing for valence information but not distinguishing modality-specific from modality-general information. In all cases, classification proceeds iteratively until each fold serves once as a test set. Classification accuracies are computed by averaging the classification accuracies across folds. If classification is successful, accuracies should be significantly higher than the chance level accuracy. For each individual, the significance of classification accuracy will be evaluated based on the binomial distribution $B(n, p)$, where n is the number of trials of each classification computation and p is the probability of correct classification when the exemplars are randomly labeled (Pereira et al., 2009). For group data, a one sample t-test can be computed on the accuracies to see if they are significantly greater than chance, indicating significant classification accuracy for the group.

Above chance accuracy levels for whole-brain classification would demonstrate that valence can be successfully discriminated from patterns of activity present

positive vs. neutral, and negative vs. neutral) based on 4 numbers of voxels (400, 600, 1000, and 3000) and the results showed that for all types of classifications, the accuracies stayed stable across 4 types of voxels.

throughout the brain. It is also expected that voxels contributing to successful classification are in prefrontal areas such as orbitofrontal cortex (Chikazoe et al., 2014; Lindquist et al., 2015) or medial prefrontal cortex (Chavez & Heatherton, 2014; Kim et al., 2015; Peelen et al., 2010), and temporal areas such as superior temporal sulcus (Peelen et al., 2010). These areas have been found to be involved in modality-general processing of affect. If modality-general processing of affect holds, then valence should be predicted above the chance level in cross-modal classification. If modality-specific processing holds, then it should be difficult to predict global affective states when trained on both modalities. Lack of cross-modal decoding, however, does not in itself demonstrate modality-specific processing as null cross-modal classification results may be due to lack of power.

2.2.2. Within-Subject ROI-based MVPA

To determine if patterns of brain activity located in specific brain regions elicited by processing concepts can be used to predict the affective states, an ROI-based MVPA was utilized. The ROIs consist of anatomical regions identified in the previous affect studies (orbitofrontal cortex, medial prefrontal cortex, and superior temporal sulcus) (Figure 2.1). The ROIs are selected by segmenting each individual's brain into Automated Anatomical Labeling (AAL) regions (Tzourio-Mazoyer et al., 2002) and using the patterns of brain activity elicited in each region separately for classification. Data division, classification, cross-validation and classification accuracy computation will be performed identically to the whole-brain MVPA with the exception of using individual ROIs for feature selection rather than voxel replicability. Above chance accuracy levels in the ROI-based classification demonstrate which brain regions contain affect information. The same logic for cross-modal MVPA can also be applied for ROI MVPA. Again, if modality-general processing of affect holds, then valence should be predicted above the chance level in cross-modal classification.

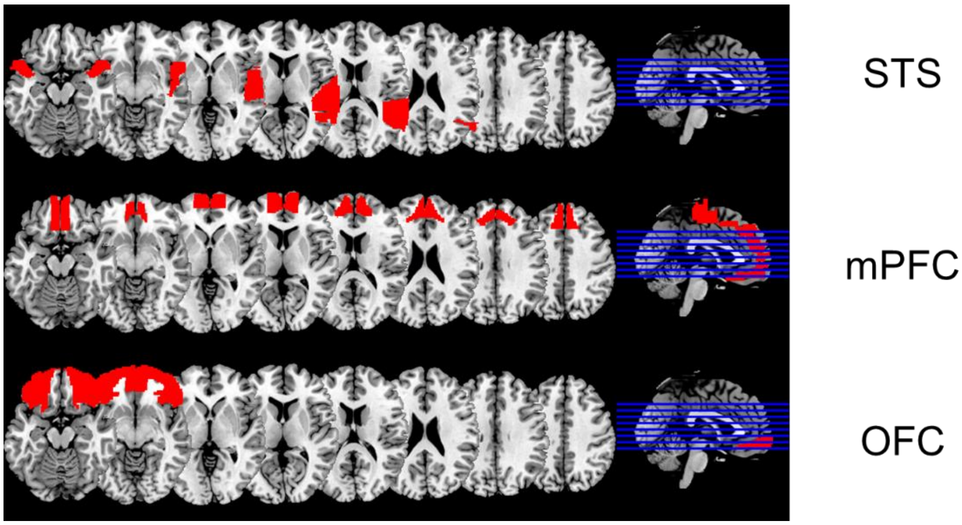

Figure 2.1 Anatomical regions of interest masks from AAL are overlaid on brain template image. Superior temporal cortex (top), medial prefrontal cortex (middle), and orbitofrontal cortex (bottom). Slice number are 56, 64, 72, 80, 88, 96, 104, and 112.

2.2.3. Cross-Participant MVPA

In order to establish commonalities between participants' neural representations of affective states, cross-participant classification will be conducted. Data from all but one participant is used to train a classifier to distinguish between affective categories. The classifier is tested on the data of the participant not included for training classifier. Classification is repeated iteratively until each participant's data serves once as a test set. Above chance accuracy levels for cross-participant classification would indicate that participants represent affective states in similar ways despite wide variation in the functional organization of human brains. The significance of classification accuracy will be evaluated based on the binomial distribution.

2.2.4. Searchlight analysis

A disadvantage of ROI-based classification methods is that they provide a spatially biased estimate of the information contained in regional patterns of brain activity, as they limit analysis to predefined anatomical regions. Rather than using anatomically defined ROIs, searchlight analysis, as an exploratory technique, finds potential regions where processing follows a particular pattern. This technique employs a spherical or cubic multivariate searchlight with a predefined search radius to scan an entire volume (Kriegeskorte, Goebel, & Bandettini, 2006). For example, a spherical searchlight radius of 9 mm was used by Connolly et al. (2010). Given cube or sphere, Connoly and colleagues (2010) applied the MVPA procedure described above and visualized regions displaying above chance classification accuracy to determine which areas of the brain contained affect information. The clusters identified by searchlight analysis should then be further examined jointly with MVPA (Connolly et al., 2012) or MDS (Kim et al., 2015) to determine whether information about the variables of interest is carried within the searchlight region. Note that a correlational approach can be applied to searchlight analysis. Rather than conducting classification in a given cube or sphere, the correlation between one modality and the other modality of voxel activation pattern can be computed. This correlation coefficient is z-transformed and resulting z maps of all participants are submitted to random effect group analysis.

Though searchlight analysis is a good analytical method to pinpoint the informative clusters in the brain, Etzel et al., (2013) listed some of limitations of this technique. For example, the size and shape of the searchlight may affect the result, and location of highly informative voxel may cause multiple voxels, which are less-informative or not-informative, to be marked as informative. This may lead to inconsistent results between searchlight analysis and MVPA. Thus they suggest validating the clusters identified by searchlight with other techniques, like classification or multidimensional scaling.

2.2.5. Multidimensional Scaling

Given the ability to successfully classify affective states, multidimensional scaling is used to investigate the lower dimensional representation from functional patterns of whole-brain activity. MDS is a statistical technique to visualize the relationship between items on a lower dimensional space. The distances in the result reflect the similarity between items. Oftentimes the interpretation of the resulting dimensions generated by MDS is important. It is well known that the two primary dimensions of affect studies are valence and arousal (Russell, 1980). In the current study, the second core dimension, arousal, is controlled, thus it is not expected to have arousal dimension. However, since two modalities are included in the study, visual versus auditory stimuli are expected to be separated by MDS.

Because MDS can be applied to any kind of distances or similarities, the correlation coefficients between experimental conditions will be used for proximity measure in the study. The conditions-by-voxel mean PSC matrices for each individual were averaged across repetitions. We hypothesize that the resulting configuration of exemplars will show a clear separation between visual and auditory exemplars and negative, neutral, and positive exemplars. A modality-general representation predicts that differences in affect are well captured by a two dimensional model in which valence differences among videos and music are captured on one dimension and modality differences are captured on a second dimension (Figure 2.2A). A modality-specific representation might include additional dimensions in which valence differences are captured for one modality but not the other (Figure 2.2B). Both modality-general and modality-specific representation could be found in a four dimensional space (Figure 2.2C). To evaluate these hypotheses, solutions will be rotated to the design matrix for these effects. Although MDS does not include statistical tests, ANOVAs can be run on dimensional values to test for hypothesized effects of valence (although these tests may not be particularly powerful if the number of exemplars is small).

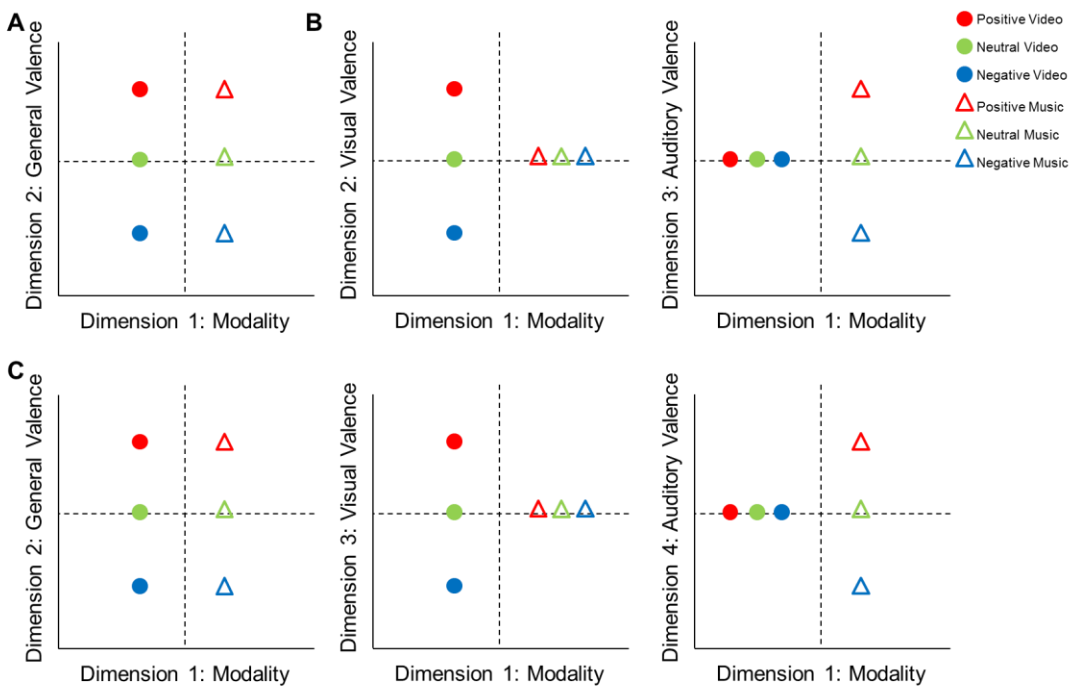

Figure 2.2 Theory-based representations of the affective space for positive, neutral, and negative valence generated by video and music. (A) Modality-general valence in a two-dimensional space. (B) Modality-specific valence in a three-dimensional space. (C) Modality-general and modality-specific valence in a four dimensional space.

2.2.6. Representational Similarity Analysis

Representational similarity analysis (RSA) is a technique that examines the representational dissimilarity between patterns of brain activity to compare how the information content is represented (Kriegeskorte, Mur, & Bandettini, 2008). One application of this technique is to compare the structure of information carried in the brain patterns elicited by two conditions across different brain regions. This technique is unique compared to MVPA, searchlight, and MDS because brain activity from whole brain or ROIs can be directly compared to each other or to multiple conceptual models. A general procedure of RSA is shown in Figure 2.3. For each brain region, the conditions-by-voxel mean PSC matrices for each individual were included in the analysis. At the same time, conceptual models such as valence only model, modality only model, or combined model, are constructed by computing the Euclidean distances between the exemplars from design matrices. Dissimilarity matrices are created by subtracting the computed correlations from 1, where a value of 1 reflects no correlation, 0 reflects a perfect correlation, and 2 reflects a perfect anti-correlation. The dissimilarity matrices are vectorized and combined. A set of nonlinear multiple regression will be performed using absolute values of coefficients to evaluate what theoretical dimensions should be included using standard regression model testing.

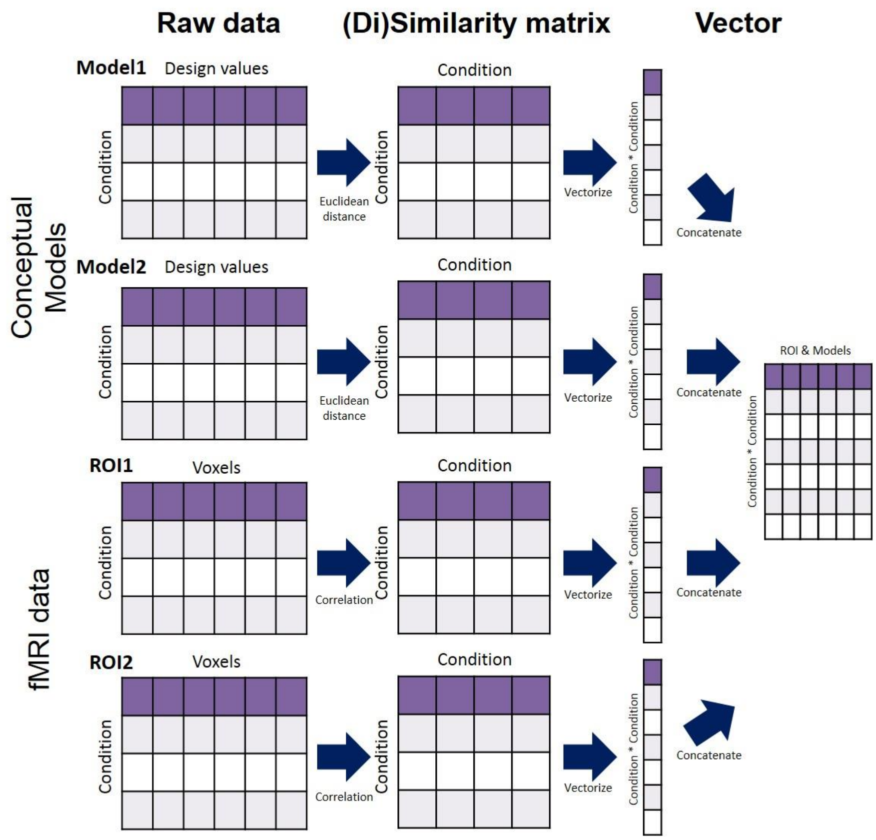

Figure 2.3 General procedure of representational similarity analysis. The format of the raw data is either condition-by-voxel matrix of fMRI-derived values or a matrix of design values. For each type of data, proximity matrix was generated by computing either Euclidean distance matrix or a matrix of correlation coefficients. Proximity matrices for each region were vectorized and concatenated.

2.3. Goals of the current study

The primary goal of this research is to explore the nature of valence representations in the brain. This will be tested at the whole brain level, in predefined ROIs, and in areas uncovered by searchlight. A key test in each of these brain regions is to see if cross-modal classification is possible, as this type of classification requires modality-general processing of affect. Evaluation of modality-specific affective processing in the brain will be considered, especially as it relates to sensory areas of the brain and the type of attentional focus required by the task. Various methods will allow us to test the hypothesis that focusing attention on the affective aspect of stimuli will enhance modality-general processing, whereas focusing on the non-affective aspect of stimuli will result in mostly modality-specific processing.

This research uses two different stimulus modalities, silent videos and music. Previous studies have utilized static visual pictures or auditory everyday sounds (i.e. Baucom et al, 2012; Shinkareva et al., 2014) or controlled dynamic stimuli, such as vocalizations, and changing body and facial expressions (i.e. Peelen et al., 2010). The current study will utilize naturalistic dynamic video stimuli which provide better externally validity. Advantages of using silent videos and music are that both unfold over time, both can elicit strong affective responses, and this manipulation clearly isolates different stimulus modalities. Shinkareva et al. (2014) presented two stimuli of the same valence and modality back to back within each trial so that a presentation of two picture stimuli had a different timing from a presentation of two sound stimuli (for the auditory trial, there was a temporal change within each sound presentation whereas for the visual trial, there was only one temporal change between two picture presentations). Thus the modality-specific processing they found might be due to the inconsistent timings between the two modalities. Silent video and music in the current study will have similar timings.

For each mode, stimuli will be drawn from one of three valence categories, negative, neutral and positive and matched on arousal level. A recent meta-analysis study (Lindquist et al., 2015) tested three valence hypotheses (bipolarity, bivalent, and general-valence) and found the most evidence for the general valence hypothesis and some evidence for the bivalent hypothesis (separate systems for negative and positive affect encoding). According to the general valence hypothesis, voxels sensitive to affective processing may show little difference between positive and negative stimuli but may differentiate these from neutral stimuli. By utilizing stimuli rated positive, negative, and neutral, I will be able to compare three valence categories. Thus, the analyses are able to consider issues of how valence is represented as well as consider affect more broadly than if only positive and negative stimuli were used. The stimulus design may be particularly helpful in that the valence categories are matched on arousal. Because in other studies, neutral stimuli, the support for general-valence may be due to a confounding relationship between arousal and valence. This was not the case in the current study.

The specific aims of this study are as follows:

1. **Determine if affective information is detectable from the patterns of brain activity elicited by processing of affect.** First, we must establish that affective states are represented in the fMRI data. The hypothesis is that the whole brain activation pattern will provide information for decoding affective states in each of the stimulus modalities, visual and auditory (silent video and music stimuli). Specifically, I hypothesize that all three valence categories will be identified from the brain activation pattern from the whole brain or a priori-defined regions of interest. This is tested with within-participant MVPA. Demonstrating classification accuracy is significantly higher than chance would support the conclusion that valence can be identified from the whole brain activation pattern. All three two-way classifications (positive vs. negative, positive vs. neutral, and negative vs. neutral) will be performed separately. Once the valence information is found, then modality-general processing of affect will be tested with cross-modal classifications at the whole brain level and predefined ROIs. Searchlight analysis will be used to identify clusters that are

involved in modality-general processing of affect and those clusters will be validated with additional confirmatory analyses in order to demonstrate if the voxels in the clusters are informative of modality-general valence processing. I also hypothesize that the structure of the internal representation of valence is similar across individuals, which will be tested with cross-participant classification. Collectively, success with multi-voxel pattern analyses would demonstrate that patterns of brain activity contain information pertaining to affective states.

2. **Determine if modality-specificity depends on the attentional focus.** The second set of working hypotheses includes that intentional evaluation of affective aspect of the stimuli will result in modality-general representation of valence, whereas focusing on semantic or perceptual features of affective stimuli will result in modality-specific representation of valence. It is hypothesized that when participants focus on affective aspects of the stimuli, then not only should affect be decoded within a given modality but also that cross-modality decoding of affect should occur. On the other hand, when participants focus on non-affective aspect of the stimuli, then decoding should be successful within-modality, but not across-modalities. This is tested by comparing the number of subjects with significant classification accuracies.

   Also I plan to examine the lower dimensional representations of the observed neural patterns of whole brain activity under different experimental conditions. Specifically, it is hypothesized that the lower dimensional representation from the affect-focusing condition will show the modality-general valence dimension which separates both visual valence and auditory valence, whereas the lower dimensional representation from the non-affect focusing condition will show a visual valence dimension, which distinguishes positive, neutral and negative visual valences but not auditory valences, and the auditory valence dimension, which distinguish positive, neutral and negative auditory valences but not visual valences.

One issue to consider is the nature of affect processing in areas of the brain that are considered modality-specific. We hypothesize that processing in these areas will be modality-specific as revealed in classification and MDS analyses. However, we will test for the degree to which valence from the incongruent modality is processed in a modality-specific region, such as representing music affect in a visual region or video affect in an auditory region. There is recent evidence for this type of multi-modal mapping within a modality-specific region (Meyer et al., 2010).

CHAPTER 3

STIMULUS SELECTION

3.1. Purpose

The purpose of the behavioral experiments was to develop and validate the stimuli to be used in the main fMRI experiment. Because I wish to explore the representation of valence for both visual and auditory stimuli while controlling for arousal, the most challenging aspects of this behavioral experiment were 1) to find stimuli which are as different as possible in terms of valence *and* as similar as possible in terms of arousal, and 2) to match the valence and arousal ratings of videos and music clips as closely as possible. Additionally, in the video clips, semantic categories (human, animal, and scene) needed to be balanced across the three valence categories (Shinkareva et al., in preparation). Finally, there was also an attempt to match low level features across valence categories. It has been reported that valence and arousal responses to stimuli are systematically related to the low level features. For example, Lakens et al. (2013) found that in emotional picture stimuli sets, the positive pictures were overall brighter than the negative pictures. For music stimuli, major mode often conveys happiness or joy, while minor mode is associated with sadness (Gabrielsson & Lindstrom, 2001; Laurier et al., 2009; Webster & Weir, 2005). Loudness is known to be positively associated with arousal (Gabrielsson & Lindstrom, 2001; Juslin & Laukka, 2004). These previous studies imply that valence representation in the current study can be confounded with the low level features. To reduce the effects, the stimuli were selected to be as similar as possible across valence categories and as similar as possible in terms of low level features. For example, positive, negative, and neutral videos were selected such that the level of motion was similar across the valence categories. After the final selection of the stimuli, the low level features were compared across the affective categories in an ANOVA. Although matching across all low level features is not possible, there was an attempt to control these as much as possible.

The behavioral experiments were conducted in order to pick ten unique exemplars for each valence and modality category while controlling for arousal and low level features (60 exemplars in total). Four independent behavioral studies were run. The general procedure was similar across the four studies. First, 20 best-guess exemplars were generated for each valence and modality category. Second, valence and arousal ratings for each stimulus were collected in a behavioral validation study. The response scale consisted of a 9 by 9 grid in which the vertical dimension was arousal (1= low, 5 = moderate, 9 = high) and the horizontal dimension represented valence (1 = negative, 5 = neutral, 9 = positive, with corresponding shading and color as shown in Figure 3.1 (Russell et al., 1989). To exclude stimuli with high and low arousal ratings, only stimuli between 4 and 6.5 points on arousal dimension were selected. Additional stimuli were generated as needed, and this procedure was repeated four times until a final set satisfied these requirements.

3.2 Study 1 through 4

3.2.1. Method

Participants

There were a total of 84 participants across studies (37, 14, 13, and 20 in each of the 4 studies). All participants were recruited from the University of South Carolina Psychology Participant Pool. Informed consent was obtained from each participant prior to the experiment, in accordance with the protocol set forth by the University of South Carolina Institutional Review Board. All participants were naive with respect to the hypotheses under investigation.

Stimuli

Participants viewed affect inducing video clips or listened to musical clips. Video stimuli were collected from internet sources (Youtube, https://www.youtube.com/, and Vimeo, https://vimeo.com). Music stimuli were obtained from Eerola and Vuoskoski's (2011) original soundtrack or an internet source. The music clips were primarily orchestral and devoid of vocals, rhythmic, and electronic instrumentation. Each excerpt was 4 s in duration. Twenty exemplars were collected for each modality and each valence category.

Procedure

The general time course of the experiment is depicted in Figure 3.1. A blank screen was presented for 500ms, followed by a 4 s of stimulus presentation. After each stimulus presentation, participants were asked to indicate their affective response to the stimuli along one of two dimensions reflecting the degree to which the participant reported feeling valence (negative to positive) and arousal (low to high). No time constraints were given for the ratings. The order of presentation of the pictures and music trials was random for each participant.

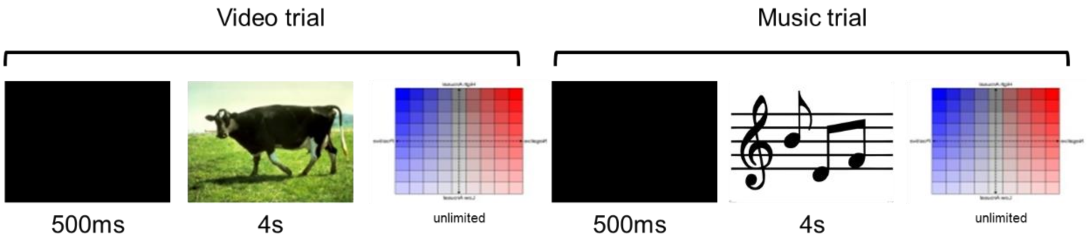

Figure 3.1 A schematic representation of the presentation timing. A single video and a single music trial are shown.

3.2.2. Results

Analysis of behavioral ratings

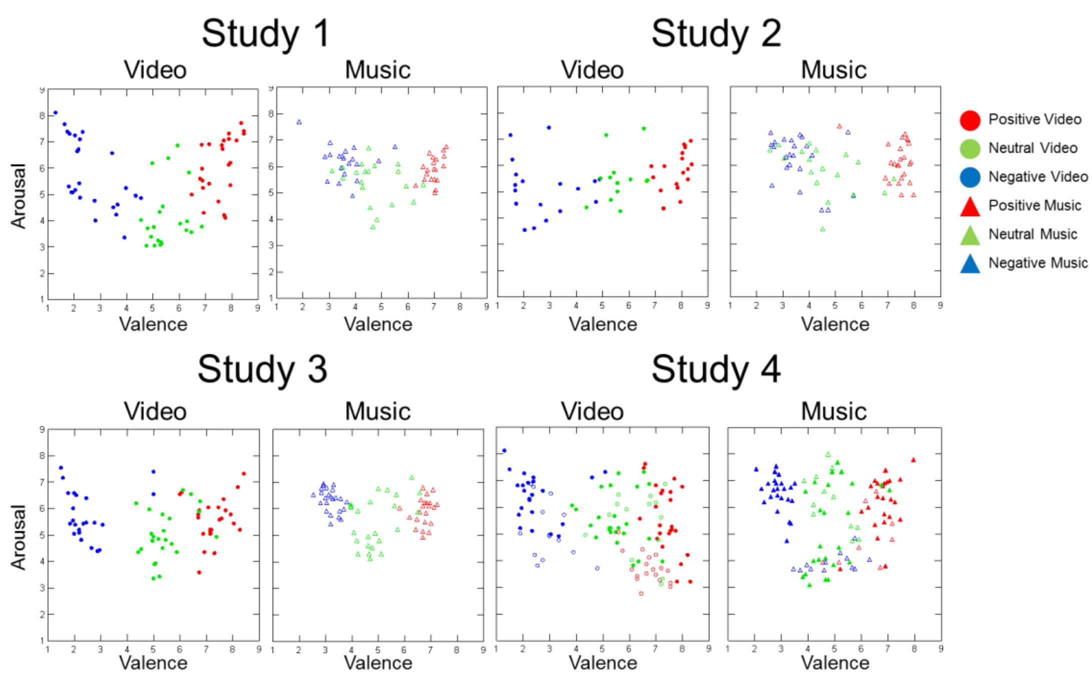

Figure 3.2 Stimuli varied on valence and arousal. Average ratings across participants shown.

Figure 3.2 illustrates the results of the 4 stimuli validation studies. The horizontal dimension represents the valence rating, and the vertical dimension represents the arousal rating. Only the exemplars between 4 and 6.5 in arousal ratings were selected for the next study. New stimuli were added to the follow up study to make the total number of exemplars over twenty.

All procedures were the same for the four studies with the following exception for Study 4. An additional replicate was created for each exemplar of video and music stimulus. For example, suppose that the 'dancing with daughter' exemplar was cropped from the original video clip and has survived the first three studies. Then for Study 4, the additional 4 s clip was cropped from the original video clip from a different time window. The final stimuli sets are shown in Figure 3.3.

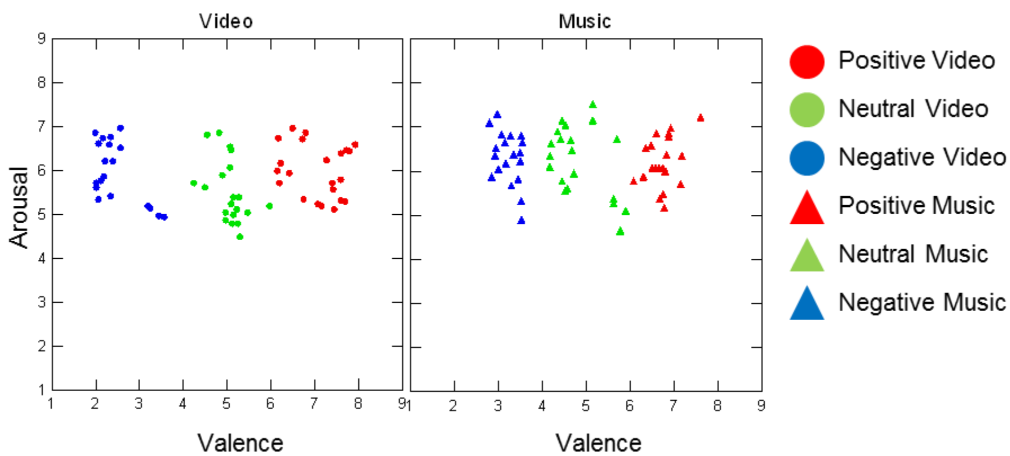

Figure 3.3 The final set of ten exemplars consisting of two replicates each.

Separate ANOVAs evaluated the valence and arousal ratings as a function of the 3×2 (valence × modality) design. The ANOVA on valence rating revealed a significant main effect of valence, $F(2,116) = 154.41$, $p < .001$, with specific comparisons indicating that positive stimuli were rated more positively than neutral and negative stimuli, and negative stimuli were rated significantly more negatively than neutral and positive and neutral stimuli. The main effect of the modality was also significant, $F(1, 116) = 5.02$, $p < .05$, indicating that videos were more positive than music clips. The interaction between modality and valence was also significant, $F(2,116) = 154.41$, $p < .001$, with simple effects analyses indicating that positive and neutral videos were rated more positively than the corresponding positive and neutral music clips, whereas negative music clips were rated more positively than negative videos. The ANOVA conducted on arousal ratings indicated that valence and interaction effects were not significant, *ps* > .05, whereas modality was significant, $F(1,116) = 11.74$, $p < .01$; musical clips were more arousing than videos. In sum, ten exemplars with two replicates for each valence and modality category were generated such that they were significantly different in valence ratings yet not significantly different in arousal ratings.

Analysis of low level features

Low level features of the stimuli represent possible confounding variables with valence. Because video stimuli are multimodal and dynamic, various types of low level features were measured. First, visual features RGB (red, green, and blue) or HSV (hue, saturation and value) color models are the most widely used cylindrical-coordinate representations of colors. Specifically, hue is the attribute of a visual sensation according to which an area appears to be similar to one of the perceived colors, saturation is the colorfulness of a stimulus relative to its own brightness, and value is a representation of the perceived luminance in relation to the saturation (Fairchild, 2005). MATLAB (R2010b, MathWorks) was used to read videos, and RGB matrix for each frame in each video stimulus was saved. Next, the RGB matrices were converted to HSV matrices. Mean hue, saturation, and value of each frame were computed by first averaging all pixels' HSVs and then averaged HSVs across frames to calculate the mean HSV for each video. Second, total motion parameters for each video were estimated based on absolute differences between frames, without respect to the direction of motion that causes differences. Motion estimation was performed at several time differences, from slower or drifting motions to fast transient motions. The number of differences above seven thresholds for each video stimulus were recorded.

For the music stimuli, frequency and amplitude were measured with MATLAB *MIR* toolbox (Lartillot et al., 2008). For each stimulus, mean pitch was computed using the *mirpitch* function in MIR toolbox and the global energy of the signal was computed by taking the root average of the square of the amplitude using the *mirrms* function. Finally, the tempo was estimated by detecting periodicities from the onset detection curve, using the *mirtempo* function.

Each low level feature was evaluated separately in a 3×2 (valence × modality) design. ANOVAs revealed that none of the main effects or interactions between valence and modality were significant for any of the features, $ps > .05$. These results suggest that these particular lower level features of the stimuli are not confounded with valence.

3.3. Summary

The purpose of the series of behavioral stimuli validation studies was to select and validate video and music stimuli for the main fMRI experiment. Sixty unique exemplars with two replicates were chosen for the main fMRI experiment. Valence ratings were significantly different between valence categories while valence categories were equated on arousal ratings. One of the advantages of the developed stimuli set is that the stimuli are equal in the arousal ratings. Many emotional stimuli databases such as IAPS or IADS show U-shaped distribution of valence and arousal, suggesting that positive and negative stimuli are higher in arousal compared to neutral ones, making it difficult to attribute the finding solely to valence. Critically, the valence categories in the developed stimuli set are controlled for the arousal, thus allowing us to attribute any difference between valenced and non-valenced conditions to valence differences rather than differences in arousal. A second advantage of the developed stimuli set is that the many low level features of the stimuli are also equal across valence categories. For example, if positive and negative videos are successfully identified on fMRI data, then we will have more confidence attributing those differences to valence, but not because of the brightness of the videos.

One limitation of this stimulus set is that the valence distributions of two modalities differed slightly, as evidenced by a significant interaction between modality and valence categories. Positive and neutral videos are more positive than positive and neutral music clips, whereas negative videos are more negative than negative music, indicating that the range of valence ratings of video stimuli is wider than that for music stimuli. This discrepancy between the two distributions may lead to, for example, a more successful valence classification of videos compared to music clips, which means that the difference between the two valence classifications may be due to the differences in valence distributions and not due to modality. However, it is important to note that valence rating differences were very large for both types of modalities, predictive of successful classification of valence categories for both modalities.

CHAPTER 4

BEHAVIORAL EXPERIMENT

4.1. Purpose

The purpose of the current Behavioral Experiment is to examine the characteristics of the affective and semantic tasks to be used in the fMRI experiment. One purpose of the fMRI study is to compare the representation of affective states under different attentional focus conditions. To consider to what degree differential effort of the tasks may confound the interpretation of the attentional manipulation, one of the main analyses is to compare the classification accuracies of valence under the two tasks. If the task difficulty interacts with valence categories across tasks, then classification may reflect task difficulty as well as the valence differences. This behavioral experiment was designed to examine the similarity of task related difficulties across valence categories for both tasks.

4.2. Method

Participants

Participants were 22 adults recruited from the University of South Carolina Psychology Participant Pool. Informed consent was obtained from each participant prior to the experiment, in accordance with the protocol set forth by the University of South Carolina Institutional Review Board. *Stimuli*

For the affective state induction, participants viewed video clips or listened to musical clips. These stimuli are described in Chapter 3.

Procedure

The general time course of the experiment is depicted in Figure 4.1. A blank screen was presented for 500ms, followed by 4 s of stimulus. After each stimulus presentation, participants were asked to respond to either an affective or semantic task. Specifically, the affective task requires participants to determine if the video clip or musical piece is positive or not, negative or not, or neutral or not (3 types of questions). The number of each type of affective task question (“Positive?”, “Negative?”, and “Neutral?”) was balanced across the three valence categories. In the semantic task, participants decided if the video clip was best characterized as having human, animal, or scene information or if the music clip was best characterized as having a string, wind, or percussion instrument. The numbers of the three types of semantic task questions were also balanced across the three valence categories for each modality. All of the task trials consisted of binary choices and the ratio of ‘yes’ or ‘no’ responses was balanced across the three valence categories. Participants used their index finger to respond ‘yes’ and the middle finger to respond ‘no’ on the keyboard. This mapping was consistent with responding in the scanner. The task was designed to shift attention to either affective or semantic aspects of the stimulus. Hence stimuli were presented in blocks in which the same modality and task occurred. At the beginning of the session, an instruction screen always showed the experimental condition of the current session. For example, at the beginning of the ‘Music-Semantic’ session, an instruction screen displayed: “This session is Music Emotion. All of the trials are music and all of the questions are about emotion. The questions are "Positive?", "Negative?" or "Neutral?"” (see Appendix A. for instructions for all types of the sessions).

After a stimulus presentation, one of three affect questions or one of three semantic questions was displayed. At this screen participants were asked to respond 'yes' or 'no' within 1.5 s. This task screen did not terminate when the participant responded but remained for the full 1.5 s. The presentation sequence was block randomized with the restriction that no affect by-modality condition was presented twice in a row.

Practice trials were also presented before the main trials. The practice session consisted of 12 trials (2 modalities × 3 valences × 2 tasks × 1 exemplar) with feedback or without feedback, respectively (24 trials in total). The stimuli used for the practice session were different from those used in the main experiment.

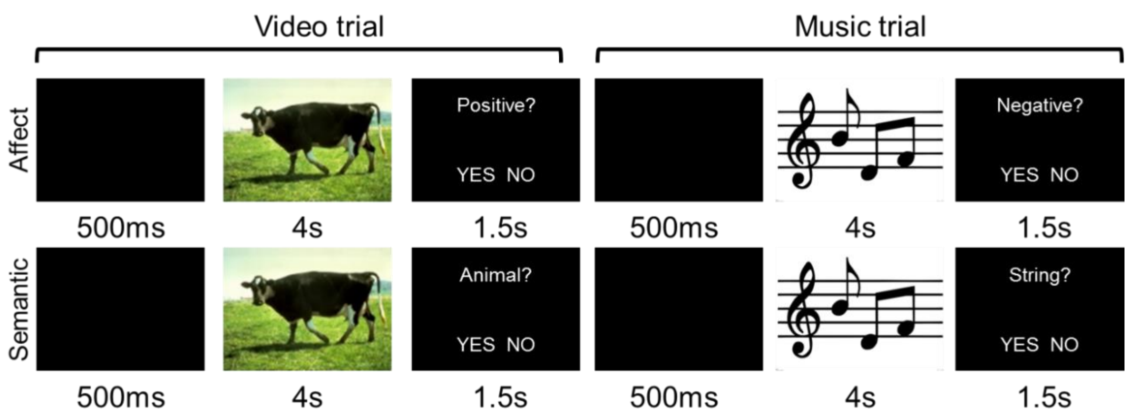

Figure 4.1 A schematic representation of the presentation timing showing affective task (top) and semantic task (bottom) of the video (left) and music (right) trials.

4.3. Results

First, 21 of 22 participants responded to over 90% of all trials (M = 98.8%). The minimum of response rate was 89.3%, and this participant was eliminated from the analyses for not meeting the criteria of a 90% response rate.

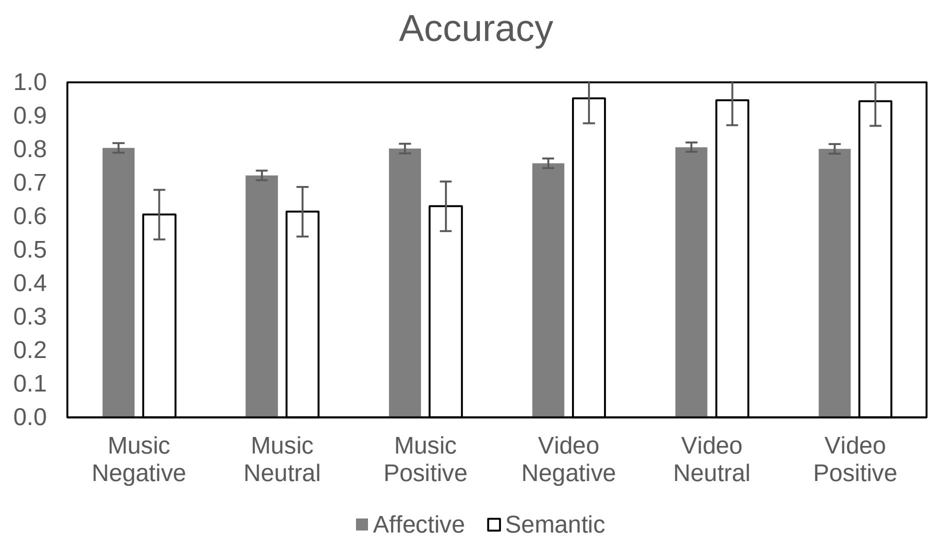

Figure 4.2 Mean accuracy of modality by valence conditions.

Figure 4.2 illustrates the mean accuracy for each experimental condition. Separate ANOVAs evaluated the accuracy and reaction time as a function of the 2×2×3 (task × modality × valence) design. The ANOVA on accuracy revealed only one significant effect, the interaction between task and modality, $F(1,21) = 180.81$, $p < .001$. Simple effect analysis indicates that there was no difference between music and video for the affective task. However, there was a significant difference between the two modalities for the semantic task, $p < .001$. Specifically, responses were less accurate for music questions (wind, string, or percussion) than for video questions (human, animal, or scene). However, no effects related to valence categories were significant, *ps* > .05.

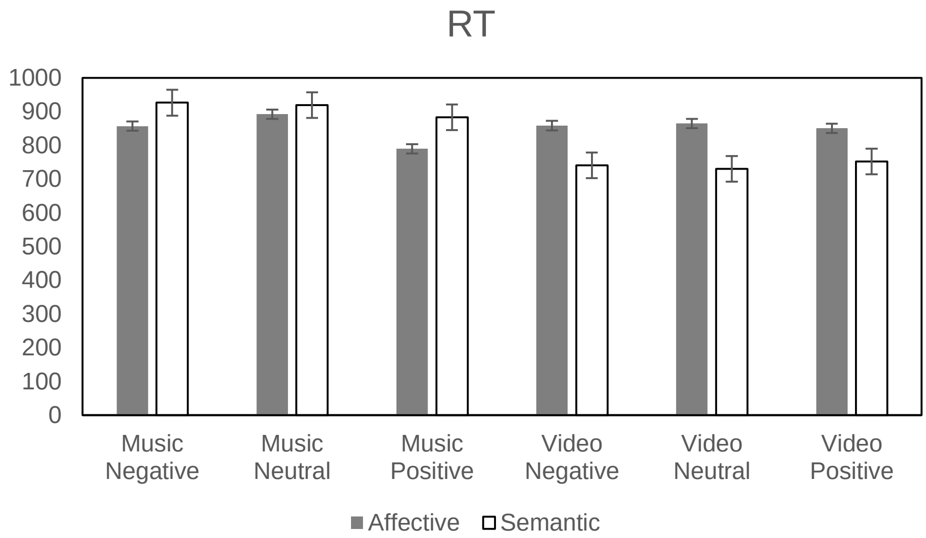

Figure 4.3 Mean reaction time of modality by valence conditions.

Figure 4.3 illustrates the mean reaction time for each experimental condition. The ANOVA again revealed a significant interaction between task and modality, $F(1,21) = 17.22$, $p < .001$, showing that there is no difference between music and video for the affective task whereas there is a significant difference between the two modalities for the semantic task, $p < .001$. Specifically, it took less time for the participants to decide if there was a human, animal, or scene in the videos than to determine if there was wind, string, or percussion in the music pieces. The interaction between valence and task was also significant, $F(1,21) = 4.81$, $p < .05$, indicating that there is no main effect of valence for semantic trials, whereas neutral trials took longer to respond to compared to positive trials, $p < .01$.

4.4. Summary

The purpose of this behavioral experiment was to explore the characteristics of the task for the main fMRI experiment and to see if there were any differences in error rates and reaction times between valence categories. The results indicated that the semantic task for music was more difficult than the semantic task for videos, but that the affective tasks were of equal difficulty for the two modes. Participants made more mistakes and were slower compared to the semantic task for the video trials. However, there were no significant main or interaction effects related to valence categories, suggesting that the valence representation or classification may not be due to task difficulty. The only significant difference between valence categories was the reaction time for affective trials, showing slower responses to neutral trials compared to the valenced trials. Because a key hypothesis in the fMRI experiment concerns how tasks may interact with affective processing, the general lack of valence effects in these analyses is reassuring. The one observed interaction effect is only linked to neutral stimuli in the semantic task, and so has limited scope.

CHAPTER 5

FUNCTIONAL MRI EXPERIMENT

5.1. Introduction

The purpose of the fMRI experiment was to explore the representation of affective states in the brain. Specifically, the goal was to determine how the affective representations elicited by stimuli differ as a function of the stimulus modalities and the attentional focus. The main hypothesis was that the neural representation of affective states while focusing on the affective aspect of the stimulus would be modality-general, whereas the neural representation of affective states while focusing on the non-affective aspect of the stimulus would be modality-specific. The analyses were aimed at two main objectives. The first was to test for modality-general and modality-specific representations of valence at the whole brain level and determine under what conditions these representations occurred. The second was to test for modality-general and modality-specific representations for different brain areas, which were either predetermined ROIs or clusters emerging from searchlight analyses. Also, because neutral stimuli were included, additional analyses were directed toward testing theories of how affect is represented in the brain according to one of three models: bipolar, bivalence and general valence.

5.2. Method

Participants

Twenty healthy volunteers (6 males) ranging in age from 20 to 32 years (M = 24) with no history of neurological disorders participated in the study after having given written informed consent. Experimental procedures were approved by the University of South Carolina Institutional Review Board (Pro00042480). All participants were right-handed, native speakers, and naive with respect to the hypotheses under investigation. The basic design was based on the factorial combination of valence (positive, neutral, and negative), stimulus modality (visual and auditory), and task type (affective and semantic), all manipulated within-participants. Dependent variables included the neural responses and judgments of affective dimensions described below.

Stimuli

The silent video and music stimuli represented two modalities (visual or auditory) and were drawn from three valence categories (positive, neutral, and negative). The procedure to develop the stimulus sets is described in the previous chapter.

Procedure

The study was composed of three sessions: the practice session, the main experiment in the MRI scanner, and the follow-up behavioral task session. During the functional image acquisition, participants performed three blocks of 120 trials of an affective or semantic evaluation task. Specifically, the affective task required participants to determine if the video clip or musical piece was positive or not, negative or not, or neutral or not. The numbers of the three types of affective comparisons were balanced across three valence categories. In the semantic task, the participants decided if the video clip was best characterized as having a human, an animal, or just a scene for video stimuli and determined if string, wind, or percussion instruments were played in the music clip. Each video clip had only one out of three elements whereas a music clip might have more than one instrument. Thus, each video clip had only one 'yes' response as the correct answer (i.e. if one video had 'human', then the correct answer to 'Human?' is 'yes' whereas the correct answers to 'Animal?' and 'Scene?" were 'no'), whereas each music clip might have multiple 'yes' responses as the correct answer (i.e. if one music clip had 'wind' and 'percussion', then the correct answer to 'String?' is 'no' whereas the correct answers to 'Wind?' and 'Percussion?" were 'yes'). Again, the numbers of the three types of semantic comparisons were balanced across three valence categories for each stimulus type. All of the task trials were binary choices and the ratio of responding 'yes' or 'no' was balanced across three valence categories (1/3 'yes' in both the affective and semantic task). Participants always used their index finger to indicate 'yes' and middle finger to indicate 'no' in the scanner. This type of task was designed to make the participant focus on the affective aspect or semantic aspect of the stimulus. The initial fixation cross was presented for 350 ms followed by 4000 ms music piece or video clip and a 5000 ms fixation cross. Music stimuli were delivered via Serene Sound Audio System (Resonance Technology, Northridge, CA). After the fixation cross, participants were asked to choose between "yes" or "no" given the affective or semantic comparison with a response deadline of 1500 ms (Figure 5.1). This task screen did not terminate when the participant clicked a button but only after the response deadline. The presentation sequences were block randomized with the restriction that no affect by-modality condition was presented twice in a row.

Before scanning, each participant performed practice trials outside of the scanner in order to get familiar with performing the task in the scanner. Examples of each semantic category (human, animal, and scene of video stimuli and wind, string, and percussion of music stimuli) and each valence category (positive, neutral, and negative for both modalities) were presented first. The practice trials were designed to be exactly the same as the procedure of performing the task in the scanner. The practice session consisted of 12 trials (2 modalities × 3 valences × 2 tasks × 1 exemplar) with feedback and without feedback, respectively. Altogether 24 practice trials were presented. The stimuli for the practice session were not from the actual stimuli set.

Functional MRI data were acquired in six runs. Within each run, 60 trials were presented in 2 blocks (affective and semantic task) of 30 trials (three valence categories and 10 unique exemplars). Each block started with an introduction screen that told participants about the modality and type of task and trials in each block. Blocks were presented in pseudorandomized order and counterbalanced across participants. Both the task and modality were blocked. One repetition consisted of 60 trials, which were three valence categories (negative, neutral, and positive), ten unique exemplars, and two types of task (affective and non-affective) for either video or music, and repeated three times. Because the estimated time duration for a single repetition was 22 minutes, one repetition was divided into two runs, blocked by modality.

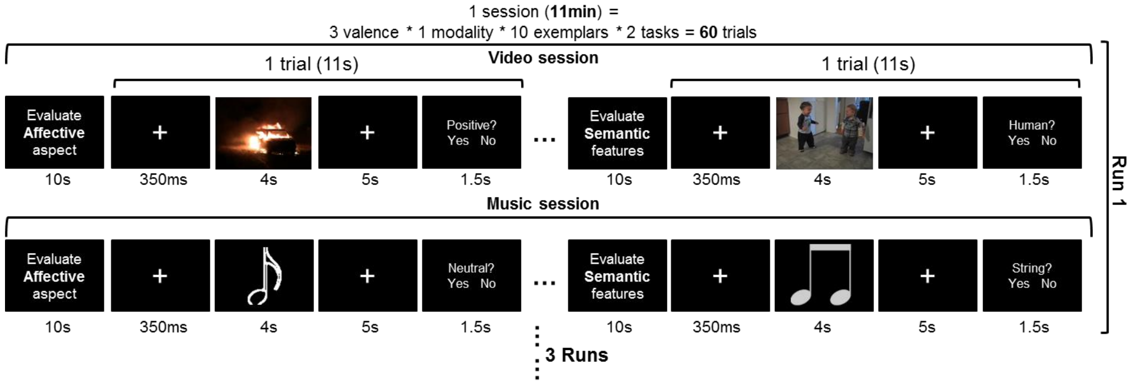

Figure 5.1 A schematic representation of the presentation timing showing video and music trials in fMRI experiment

In addition to performing the fMRI task, after scanning the participants completed a behavioral task to rate the perceived emotion intensity displayed on the affective grid (direct ratings on valence and arousal). All of 120 unique stimuli were presented using the same procedure described in Chapter 3 for preliminary studies. All of the trials in the practice session, main session in the scanner, and follow-up behavioral session were presented using E-Prime software (Psychology Software Tools, Sharpsburg, PA).

MR data acquisition and preprocessing

Magnetic resonance images of the brain were obtained using a Siemens Magnetom Trio 3.0T scanner (Siemens, Erlangen, Germany) at the McCausland Center for Brain Imaging at the University of South Carolina using a standard 16-channel head coil. Functional images were obtained using a gradient echo EPI pulse sequence: TR = 1550 ms, TE = 2.26 ms, flip angle = 9°, FOV = 256 mm. The acquisition matrix was 64×64 with 3×3×3 mm voxels. Preprocessing of the fMRI data was performed using SPM8 (Wellcome Department of Imaging Neuroscience, University College of London, London, UK). The EPI data were corrected for slice-timing difference, realigned for motion correction, coregistered to the individual T1-weighted images. A high-pass filter (.008 Hz cut off) was used to reduce low-frequency noise. Images were spatially normalized to standard Montreal Neurological Institute space. I fit the time-series data for each voxel using *GLMdenoise* to improve signal-to-noise ratio (Kay et al., 2013). The GLMdenoise used the 12 experimental conditions (3 valence × 2 modalities × 2 tasks) for regressor estimation but was blind to valence categories, and thus did not bias the results when comparing across combined categories (i.e., positive vs. negative or music vs. video). The GLMdenoise was applied to each fMRI session separately.

The data preprocessing steps and MVPA analysis employed in the current study are similar to those that have been successfully used in other MVPA studies (Baucom et al., 2012; Shinkareva et al., 2014). The percent signal change (PSC) relative to the average activity in a voxel was computed for each voxel in every volume. The mean PSC of five volumes, offset 3.1 s from the stimulus onset (to account for the delay in hemodynamic response), was used as the input for further analyses (Figure 5.2). Furthermore, the mean PSC data for each voxel was standardized to have a mean of zero and variance of one.

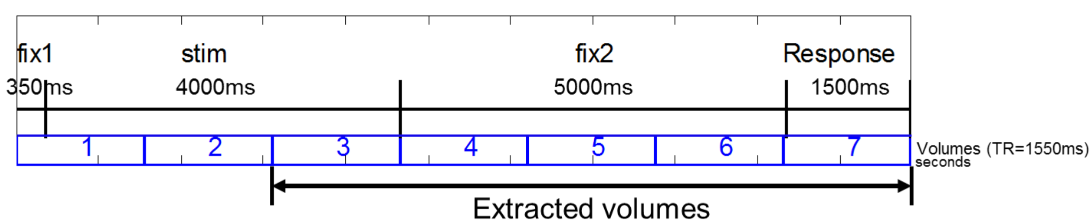

Figure 5.2 A schematic representation of the presentation timing and data extraction for a single trial.

Within-participant pattern classification

Classifiers were trained to identify affective states from the pattern of brain activity MPSC elicited by different types of the affective states across two modalities. Three two-way classifications (positive vs. negative, positive vs. neutral, and negative vs. neutral) were performed to identify relevant affective state differences. For cross validation, all trials were divided into training and test sets, and relevant voxels were extracted based on the training set only. The classifier was constructed using the selected features from the training set. The classifier was subsequently tested on the unused test set and classification performance was evaluated with multiple cross-validations. A logistic regression classifier was used for classification (Bishop, 2006). Six fold cross-validation was used to evaluate classification performance, where each fold corresponded to one block of each of the conditions. For example, all neutral trials were dropped for the positive vs. negative classification, then logistic regression classifiers were trained from 200 trials and tested to 40 trials out of 240 total trials. Classification was repeated iteratively until each presentation served as the test set once. Classification accuracies were computed based on the average classification accuracies across test folds. As a result, classification accuracy was always based upon the test data only, which remained disconnected from the training data. The significance of classification accuracy was evaluated based on 1) one sample t-test to test if the average of group accuracies is significantly higher than the chance level, and 2) the binomial distribution $B(n, p)$, where n is the number of trials of each classification computation and p is the probability of correct classification when the exemplars are randomly labeled (Pereira et al., 2009) for the individual level of analysis.

For the cross-modal MVPA, classifiers were trained from only one modality and tested on the other modality. This procedure was carried out twice so that each modality served as the test set once.

Within-Subject ROI-based MVPA

To determine if patterns of brain activity located in specific brain regions can predict the affective states, ROI-based MVPAs were utilized. The ROIs consisted of anatomical regions identified in the previous MVPA fMRI studies, including orbitofrontal cortex (Chikazoe et al., 2014), medial prefrontal cortex (Chavez & Heatherton, 2014; Kim et al., 2015; Peelen et al., 2010) and superior temporal sulcus (Lindquist et al., 2015; Peelen et al., 2010) as processing modality-general valence. The ROIs were selected by segmenting each individual's brain into Automated Anatomical Labeling (AAL) regions (Tzourio-Mazoyer et al., 2002) and by using the patterns of brain activity elicited in each region separately for classification. The procedure of data division, classification, cross-validation, and classification accuracy computation was identical to the whole-brain MVPA with the exception of using individual ROIs for feature selection rather than voxel replicability. Above-chance accuracy levels in the ROI-based classification would test the type of affect information processed during these tasks in the brain regions. The significance of classification accuracy was also evaluated based on the binomial distribution and one sample t-test.

Cross-Participant MVPA

Cross-participant classification was conducted in order to explore if there were any commonalities among individuals. Data from all but one participant were used to train a classifier to distinguish affective states associated with each experimental condition. The classifier was then tested on the data of the excluded participant. Classification was repeated iteratively until each participant's data served once as the test set. Above-chance accuracy levels for cross-participant classification would indicate that participants represent affective states in similar ways despite wide variation in the functional organization of human brains. The significance of classification accuracy was evaluated based on the binomial distribution for the individual level and a one-sample t-test for the group level.

Multidimensional Scaling

Given the ability to successfully classify affective states, multidimensional scaling was performed to investigate the lower dimensional representation from functional patterns of whole-brain activity. The conditions-by-voxel MPSC matrices for each individual were averaged across 3 presentations (three runs) of each condition to obtain 10 data points, or exemplars, for each experimental condition. Additionally, all odd and even trials were averaged to reduce noise so that there were two exemplars for each modality × valence × task condition (Baucom et al., 2012; Shinkareva et al., 2014). (Note that odd and even refer to the design matrix and not to actual trial orders.) Pairwise correlations were computed between exemplars, resulting in a 24 by 24 (2 exemplars × 2 modalities × 2 tasks × 3 valences) exemplar-by-exemplar matrix for each individual, and the correlation matrices for each individual were averaged and analyzed. A modality-general representation predicts that differences in affect are well captured by a two dimensional model in which valence differences among videos and music pieces are captured on one dimension and modality differences are captured on a second dimension. A modality-specific representation might include additional dimensions in which valence differences are captured for one modality but not the other.

Representational similarity analysis

A representational similarity analysis was applied to test neural similarity patterns against patterns predicted by different theoretical conceptual models. Here I tested three different valence hypotheses: the bipolarity hypothesis that positive and negative affect are represented monotonically, the bivalent hypothesis that positive and negative affect are supported by independent brain systems, and the general valence hypothesis that positive and negative affect are supported by valence-general regions (Lindquist et al., 2015). Table 5.1 provides the design values used to generate the pairwise distances between conditions for the 19 one-dimensional models. The $120 \times (120-1) / 2 = 7140$ distances for each type of conceptual model, and the pairwise distances were strung out in a single vector. The fMRI data sets from whole brain, anatomical ROIs, and clusters that searchlight analyses identified were extracted, and correlation was run to have condition by condition correlation matrix. Each matrix was subtracted from 1 to make the distance matrix. These distance matrices from fMRI data were strung out in single vectors. The 19 single dimension models were included in the analysis to explain fMRI data vectors and nonlinear multiple regression was applied using the absolute values of coefficients. The nature of the RSA modeling requires that model distance correlates positively with data distances, so coefficients were constrained to be positive. The original model included all terms that were significant when entered alone. In this equation, terms that were not significant were then eliminated one at a time, beginning with the term that had the smallest coefficient. The final model included only significant terms. This analysis yielded overall r square values and a change in the r square value provided model testing under the assumption that dimensions are combined in a city block space and dimensions do not need to be weighted equal.

Table 5.1. The design values used to generate the pairwise distances between conditions for the one-dimensional models. Each number represents valence hypotheses (1: bipolar hypothesis, 2: bivalent hypothesis, 3: general-valence hypothesis). Within bivalent hypothesis models, *a* models reflect 'negative over positive affect' and *b* models reflect 'positive over negative affect'. Within general-valence hypothesis models, *a* models reflect 'negative and positive over neutral affect', *b* models reflect 'relative preference for positive vs. negative affect', and *c* models reflect 'relative preference for negative vs. positive affect'.

	Design values					
Models	Video Negative	Video Neutral	Video Positive	Music Negative	Music Neutral	Music Positive
Modality (m)	1	1	1	-1	-1	-1
Video Valence 1 (vv1)	-1	0	1	0	0	0
Music Valence 1 (mv1)	0	0	0	-1	0	1
Modality-General Valence 1 (mgv1)	-1	0	1	-1	0	1
Video Valence 2a (vv2a)	2	-1	-1	0	0	0
Video Valence 2b (vv2b)	-1	-1	2	0	0	0
Music Valence 2a (mv2a)	0	0	0	2	-1	-1
Music Valence 2b (mv2b)	0	0	0	-1	-1	2
Modality-General Valence 2a (mgv2a)	2	-1	-1	2	-1	-1
Modality-General Valence 2b (mgv2a)	-1	-1	2	-1	-1	2
Video Valence 3a (vv3a)	1	-2	1	0	0	0
Video Valence 3b (vv3b)	1	-1	0	0	0	0
Video Valence 3c (vv3c)	0	-1	1	0	0	0
Music Valence 3a (mv3a)	0	0	0	1	-2	1
Music Valence 3b (mv3b)	0	0	0	1	-1	0
Music Valence 3c (mv3c)	0	0	0	0	-1	1
Modality-General Valence 3a (mgv3a)	1	-2	1	1	-2	1
Modality-General Valence 3b (mgv3b)	1	-1	0	1	-1	0
Modality-General Valence 3c (mgv3c)	0	-1	1	0	-1	1

Searchlight analysis

Two types of searchlight analyses were conducted. The first one was designed to locate the regions showing similarity patterns between the two modalities based on valence ranks. For each voxel *v*, a searchlight of a 5×5×5 cube centered on the voxels *v* was selected. Within each cube, an exemplar by voxels data matrix was extracted. Six repetitions were averaged and the resulting 60 exemplars were rank ordered based on the average of valence ratings of all participants. This data matrix was split into two: video and music trials. Each data matrix was vectorized and two vectors were correlated. The correlation coefficient reflects the similarity between the activation patterns of the valence representation of two modalities. The correlation coefficient was Fisher z transformed and assigned to the voxel *v*. This procedure was repeated through the whole brain. Each individual's z maps were submitted to a random effects whole brain group analysis to identify commonalities of the valence similarity pattern between modalities among individuals.

The second searchlight analysis was designed to locate sensory areas regardless of affective states. Specifically, similarity patterns within each modality were estimated by average of within-modality correlation coefficients. For each voxel *v*, a searchlight of a 5×5×5 cube centered on the voxels *v* was selected. Within each cube, an exemplar by voxels data matrix was extracted. Six repetitions were averaged so that 60 exemplars by voxel matrix was prepared. A Pearson correlation was performed on the transposed data matrix, then exemplar by exemplar correlation matrix was computed. Of these 30*29/2 correlation coefficients reflect within visual modality, 30*29/2 correlation coefficients reflect within auditory modality, and 30*30 correlation coefficients reflect across-modalities. All correlation coefficients were averaged for each category (within visual, within auditory, and cross-modality), and each averaged coefficient was assigned to voxel *v*. The correlation coefficient was Fisher Z transformed and assigned to the voxel *v*. Altogether, three z maps were computed for each individual. The same group analyses were run to identify modality areas. This process was similar to Peelen and colleagues (2010). Primarily, it was expected that within-modality classification would be successful. For example, the visual area will represent visual valence, and the auditory area will represent auditory valence. Additionally, it will be determined if each modality area may represent the other modality's valence. It has been reported that the subjective auditory experience evoked by visual stimuli in the absence of auditory stimulation was associated with content-specific activity in early auditory cortices (Meyer et al., 2010). If this is true, then early visual cortex, for example, may have auditory valence information and vice versa.

5.3. Results

5.3.1. Behavioral results for fMRI participants

Mean ratings on affective dimensions

The distribution of valence and arousal ratings of the stimuli are shown in Figure 5.3. ANOVAs evaluated the valence and arousal ratings separately as a function of the 3×2 (valence × modality) design. The ANOVA on valence ratings revealed a significant main effect of valence categories, $F(2,114) = 395.605$, $p < .001$. Pairwise contrasts indicated that positive stimuli were rated more positively than neutral and negative stimuli, and negative stimuli were significantly more negative than neutral stimuli. Unlike the result of behavioral studies from a separate group of participants, the main effect of modality on valence ratings was not significant, $F(1, 114) = .287$, $p > .05$, indicating that there was no significant difference between the two modalities on valence ratings. The interaction between modality and valence category was not significant, $F(2,114) = .168$, $p > .05$, showing that the distribution of valence ratings of video and music were not different.

For arousal ratings, all tests were not significant, *ps* > .05, indicating that there was no difference between valence categories and modalities on arousal ratings. These results show that ten exemplars with two replicates for each valence and modality category were used in the study such that there was a significantly difference across valence categories while equating the arousal level. One of the differences from the behavioral study of separate groups of participants was that there was no effect of modality. In the behavioral study (Chapter 3), the modality and interaction effects were significant, suggesting that there is a difference between video and music stimuli mean valence ratings and a difference between valence distributions of video and music stimuli. However, twenty participants who participated in the main fMRI study rated video and music equally on valence ratings, and the difference between valence categories in video stimuli were not different from the differences between valence categories in music stimuli. These results allow us to attribute any difference between valence distributions of two modalities to modality.

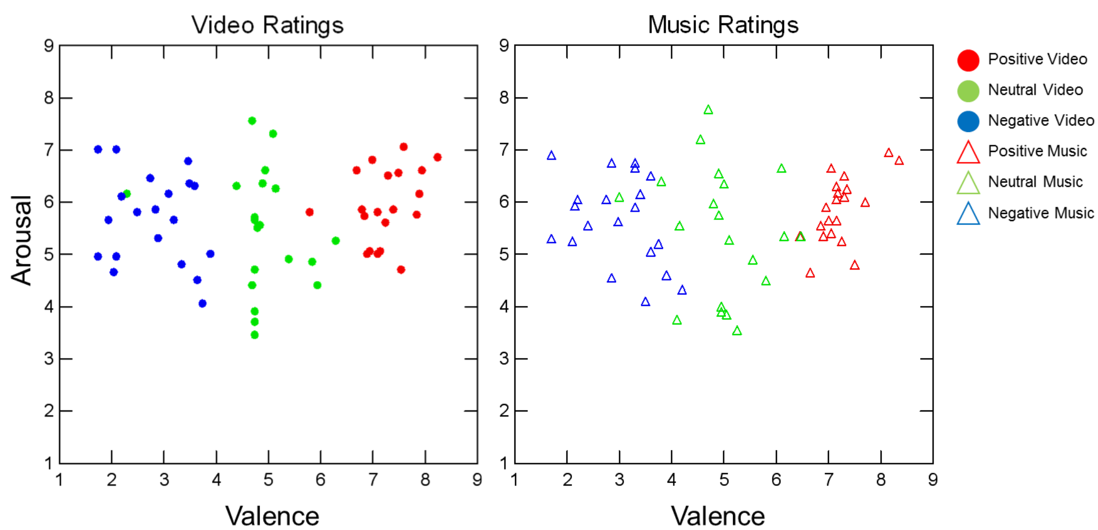

Figure 5.3 Stimuli varied on valence and arousal. Average ratings across fMRI participants were consistent with an equivalent manipulation of valence across-modalities and no differences in arousal across valence and modality conditions.

Results of accuracy and reaction time

First, the number of responded trials was investigated. 19 out of 20 participants responded to over 90% of all trials (M = 97.4%). One participant responded to only 18% of all trials so this participant's data were excluded for the behavioral analyses.

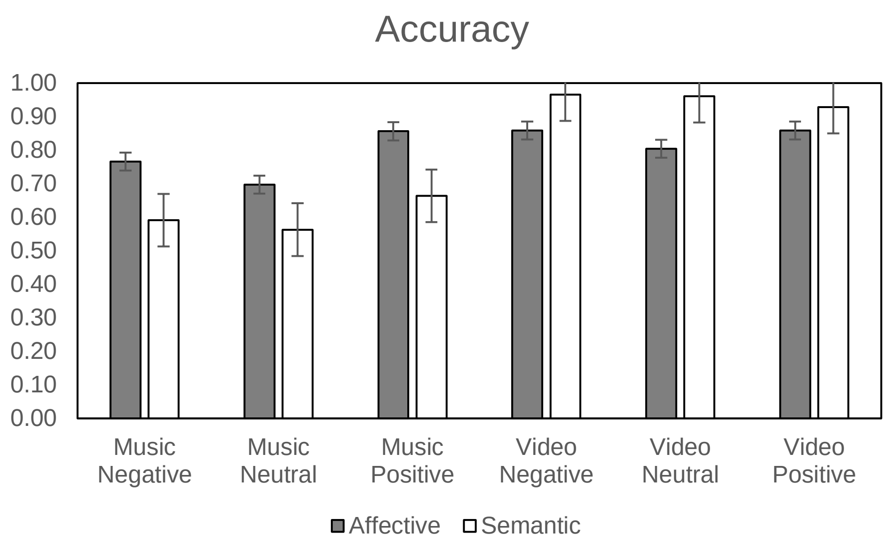

Figure 5.4 Mean accuracy of modality by valence condition

Figure 5.4 illustrates the mean accuracy for each experimental condition. Separate ANOVAs evaluated the accuracy and reaction time as a function of the 2×2×3 (task × modality × valence) design. The ANOVA on accuracy revealed a significant two-way interaction between task and modality, $F(1,18) = 261.15$, $p < .001$, and an interaction between modality and valence, $F(1,18) = 18.33$, $p < .001$. First, simple effect analysis was run for each task based on the interaction between task and modality. For the affective task, video trials were more accurate (M = .84) than music trials (M = .77), $F(1,18) = 24.34$, $p < .001$. The same relationship was found for the semantic task but the difference between the two modalities was much larger compared to the affective task, (video: M = .95, music: M = .61), $F(1,18) = 811.82$, $p < .001$. This finding is the consistent with the previous result from the separate group of participants. A simple effect analysis was conducted for each modality based on the significant interaction between modality and valence revealed that for music trials, valence was significant effect, with the highest accuracy for the positive trials (*M* = .76) followed by negative trials (*M* = .68) and then the neutral trials (*M* = .63), $F(2,36) = 18.11$, $p < .001$. For the video trials, accuracy was highest for the negative trials (*M* = .91) compared to the positive (*M* = .89) and neutral (*M* = .88) trials, $F(2,36) = 4.47$, $p = .02$. However, task type was not significantly related to valence categories, $F(2,36) = 2.62$, $p = .09$. In sum, the accuracy of the semantic trials for music trials was lower than that of semantic trials for video, and this pattern was similar to the previous result from the behavioral participants. The similar accuracy levels and patterns found for fMRI and behavioral participants support the assertion that fMRI participants were actively engaged in the task.

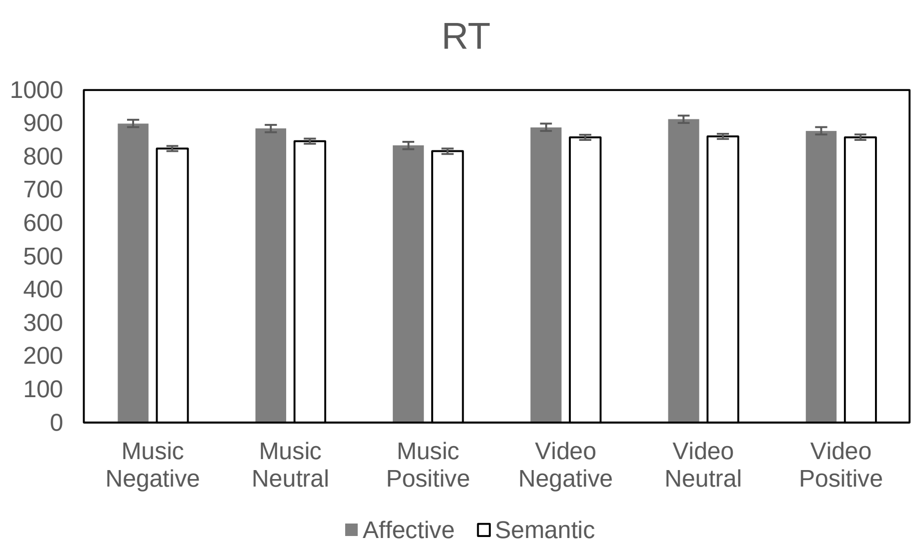

Figure 5.5 Mean reaction time of modality by valence condition.

Figure 5.5 illustrates the mean reaction time for each experimental condition. The ANOVA revealed the significant main effects of task, modality, and valence, $F(1,18) = 17.63$, $p < .001$; $F(1,18) = 4.72$, $p < .05$; $F(2,36) = 20.95$, $p < .001$. Specifically, responses to the affect task ($M = 882.1$ ms) was slower than those to the semantic task ($M = 843.4$ ms), responses to video trials ($M = 875.3$ ms) were slower than those to music trials ($M = 850.3$ ms), and responses to neutral trials were slowest ($M = 875.51$ ms) followed by negative ($M = 867.0$ ms) and positive ($M = 845.9$ ms) trials. The three way interaction was significant, $F(2,36) = 3.77$, $p < .05$, thus separate simple effect analyses were performed for each task. For the semantic task, only the main effect of modality was significant, reflecting faster responses to music ($M = 828.41$ ms) than videos ($M = 858.45$ ms). For the affective task, a main effect of valence was significant, reflecting faster responses to positive trials ($M = 855.12$ ms) than neutral and negative trials ($M = 893.33$ ms, $M = 897.97$ ms). The same RT analysis was conducted for only correct response trials and the results were very similar to the ANOVA on response trials, with the same pattern of significance.

Overall, there was no noticeable difference between valence categories for the music trials. Affective trials were slower but more accurate than semantic trials, so it is hard to conclude that affective tasks were easier than semantic task for the music trials. However, semantic trials for video were faster and more accurate than affective trials for video when the valence was positive and neutral, allowing us to conclude that video-semantic trials were easier than video-affective trials. This overall pattern is consistent with the Behavioral Study described in Chapter 4 such that the music-semantic trials were less accurate than video-semantic trials. The first difference is that in the

Behavioral Study, music-semantic trials were slower than music-affective trials but the fMRI study participants' data showed that music-semantic trials were faster than music-affective trials, so it is not possible to conclude that music-semantic trials are more difficult than music-affective trials. The second difference is that in the Behavioral Study, there were no differences between valence categories, whereas the fMRI study participants' data showed that accuracy decreased from positive to negative to neutral condition when the modality was auditory, but that the negative condition was more accurate than neutral and positive conditions when the modality was visual. These results suggest that the valence differences in classifications and MDS could be confounded with the task difficulty difference, a point that I will address more closely in Chapter 6.

5.3.2. Within-participants classification based on whole brain

First, valence classification for each modality (within-modality classification) was conducted to demonstrate if there is clear valence related information in each stimulus set at the whole brain level. Because there are three valence categories, three separate two-way classifications were performed (Positive vs. Negative, Positive vs. Neutral, and Negative vs. Neutral). One-sample t-tests were performed to evaluate if the group mean accuracies were significantly higher than the chance level (.5).

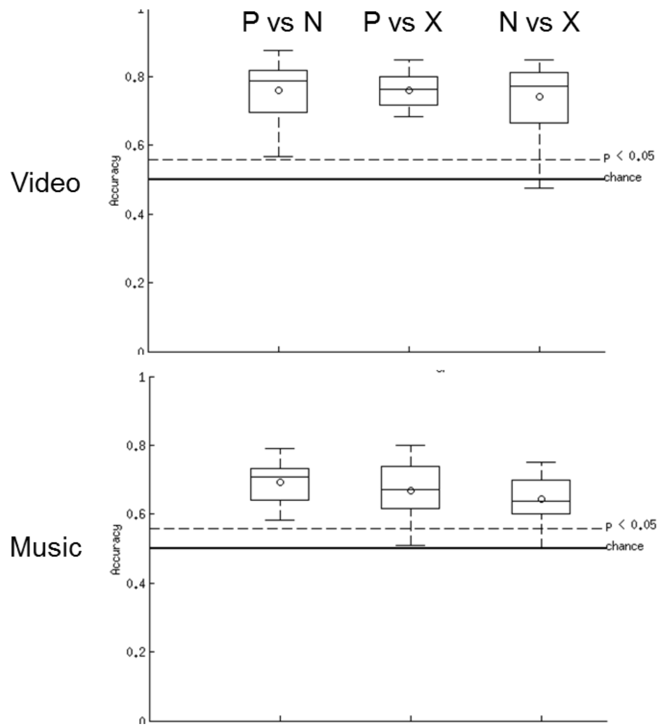

Figure 5.6 Within-participant within-modal classification accuracies for identifying trials as positive vs. negative (left box plot), positive vs. neutral (middle) and negative vs. neutral (right) for video trials (top) and music trials (bottom), summarized across twenty participants by box plots are shown.

When classifiers were trained to identify valence categories for each modality, the mean accuracies across participants were significantly greater than chance for all three types of classifications, *ps* < .001 (Figure 5.6). At the most replicable 600 voxels for each participant, classification accuracies of video for one participant were as high as 87.5%, 85%, and 85% for each classification, and accuracies for music were 78.3%, 80%, and 72.5%. Based on binomial distribution, the numbers of significant participants for videos were 20, 20, and 19 out of 20, and the numbers of significant participants for music trials were 20, 19, and 19 out of 20. This result suggests that classifiers trained from each modality and on each individual participant data were able to identify affective states of the corresponding modality's valence reliably above chance, allowing us to conclude that valence information is represented at the whole brain level when elicited by viewing silent videos or listening to musical clips within each participant. However, this does not indicate if this is modality-specific or general.

Next, within-participant combined-modality MVPA was run to test for valence ignoring the modality distinction. If this classification fails when the others succeed, then it is likely that valence specific processing is occurring. If this classification is successful, it can be due to modality-specific and/or modality-general. When classifiers were trained to identify valence categories from both modality trials, the mean accuracies across participants were significantly greater than chance for all three types of classifications, $p < .001$ (Figure 5.7). At the most replicable 600 voxels for each participant, classification accuracies for one participant were as high as 72.9%, 71.25%, and 72.5% for each classification. Based on binomial distribution, the numbers of significant participants were 19, 15, and 15 out of 20. This result suggests that classifiers trained on each individual participant data were able to identify affective states of valence reliably above chance, allowing us to conclude that valence information is represented at the whole brain level when elicited by viewing silent videos or listening to musical clips within each participant.

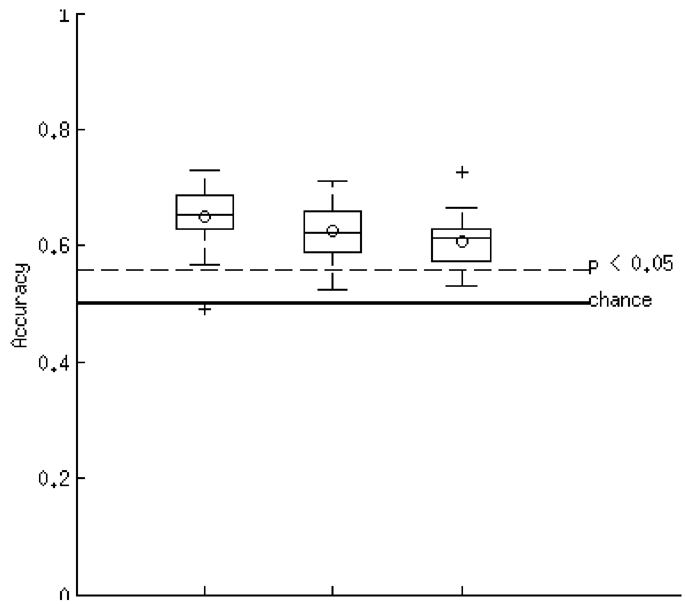

Figure 5.7 Within-participant classification accuracies for identifying trials as positive vs. negative (left box plot), positive vs. neutral (middle) and negative vs. neutral (right), summarized across twenty participants by box plots are shown.

5.3.3. Within-participant cross-modal classification

The results above indicated that valence categories could be decoded using classifiers trained from both types of the modality. To test if there is modality-general processing of affect at the whole brain level, cross-modal classification was used in which classifiers are trained from only one modality and tested on the other type of modality. Cross-modal classification is a key test of modality-general processing. If this is successful, then it validates that valence is what is being classified because there are no lower level features in common between videos and music. The procedure was that the most replicable 600 voxels were identified by correlating each exemplar over repetitions. Logistic regression classifiers were trained from one modality and tested on the other modality trials and vice versa. The results show a successful classification of valence at the whole brain level for all three of the valence classifications (Figure 5.8). The results showed that the mean accuracies were .59, .58, and .57 for positive-negative, positive-neutral, and negative-neutral classifications and one-sample t-tests revealed that these accuracies were significantly higher than chance, $p < .001$. Binomial tests revealed that accuracies of 15, 13, and 13 out of 20 participants were significantly higher than chance for each type of classification, respectively. When comparing within-participant cross-modal MVPA with within-participant MVPA, a repeated-measures ANOVA revealed that cross-modal classifications were less accurate than within-participant MVPA, $F(1,19) = 45.90$, $p < .001$. The main effect of classification type was significant, $F(1,19) = 3.65$, $p < .05$, indicating that a classification between positive and negative was more accurate than a classification between negative and neutral, $p < .05$. There was no difference between a classification between positive and negative and a classification between positive and neutral. In sum, a successful within-participant cross-modal classification result suggests that there is a modality-general processing of valence at the whole brain level across the different valence categories (positive, neutral, and negative).

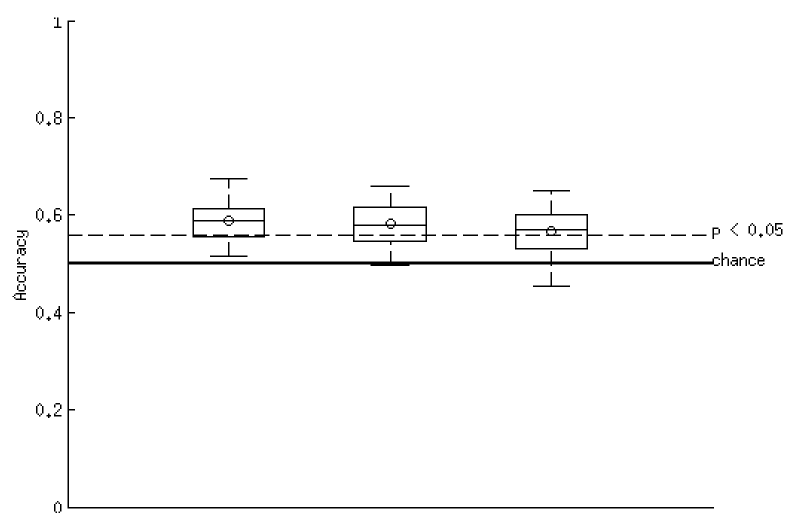

Figure 5.8 Within-participant cross-modal classification accuracies for identifying trials as positive vs. negative (left box plot), positive vs. neutral (middle) and negative vs. neutral (right), summarized across twenty participants by box plots are shown.

5.3.4. Lower dimensional representation of affective space

To examine the lower dimensional representation of affective space, the correlation matrices for odd and even trials of each of the experimental conditions were generated for each participant based on the most stable 600 voxels for each participant. The averaged 24×24 correlation matrices were then input into nonmetric MDS and a common configuration was abstracted. Because I am interested in investigating 4 a priori dimensions (modality, modality-general valence, music specific valence, and video specific valence), a 4-d solution was extracted. The final solution was transformed using a Procrustes rotation to the design matrix orientation to reflect degree of modality, modality-general valence, video- or music-specific valence. The four dimensional solution had an overall stress value of 0.21 (1d solution stress: .46, 2d: .33, 3d: .25). In the representation shown in Figure 5.9, the first dimension reflects modality, separating visual and auditory trials. The second dimension reflects modality-general valence, separating positive, neutral, and negative stimuli from both modalities. The third dimension reflects music specific valence, separating valence categories only from music trials, whereas the fourth dimension reflects video specific valence, separating valence categories only from video trials. Note that replicates of each state tended to be closer to each other in the space than to other states, reflecting reliability in classification. The only exception is neutral conditions: those tended to be clustered together or spread out between positive and negative replicates[2].

[2] The issue was that all even and odd trials for each experimental condition were averaged for clearer visualization in the current study. However, more averaging may mislead the interpretation of the MDS solution. It was argued that MDS solution from excessive averaging random data may lead to interpretations of distinctions that are not valid (Shinkareva, personal communication, April, 2015). This can be tested by taking multiple averaging approaches, such as by taking a mean of the first five and the last five exemplars. If the representation of the other type of averaging looks similar, then the

Overall, the MDS results provide a clear visualization of the data, with modality, modality-general valence, and modality-specific valence aspects of the stimuli separated in the lower dimensional space and related in expected ways to the neural responses at the whole brain level. This MDS result is consistent with the result of within-participant cross-modal classification which suggested modality-general representation of affective states. Additional ANOVAs tested for significant differences between valence categories on each of the three valence dimension's coordinates for each modality. For the dimension 2 coordinates (modality-general valence dimension), there was a marginally significant effect of valence for visual stimuli, $F(2,9) = 4.04$, $p = .055$, and a significant effect of valence for auditory stimuli, $F(2,9) = 13.56$, $p < .01$. For both of these tests, the contrast of positive and negative conditions was significant, $p < .01$. For the dimension 3 coordinates (music valence dimension), there were significant differences between valence categories only for music trials, $F(2,9) = 12.31$, $p <.01$, but not for the video trials, $p > .05$. The contrast of positive and negative conditions was significant only for music trials, $p < .01$. Finally, for the dimension 4 coordinates (video valence dimension), there was a marginally significant difference between valence categories only for video trials, $F(2,9) = 3.61$, $p = .07$, but not for the music trials, $p > .05$. Again, the contrast of positive and negative conditions was significant only for video trials, $p < .05$. Please note that this between-item ANOVA is underpowered because the number of data points for

MDS solution would be deemed more reliable. As well as averaging even and odd trials, different averaging type was conducted (the first five and the last five trials) and submitted to MDS. The resulting 4 dimensional solutions were strung out and correlated with the 4 dimensional solutions from odd and even averaging. The result showed that the two solutions based on different types of averaging are highly correlated, $r = .74$, suggesting that the MDS solution is reliable.

each modality is only 12. This set of ANOVAs and contrasts confirmed three dimensions as modality-general, music-specific, and video-specific valence dimensions.

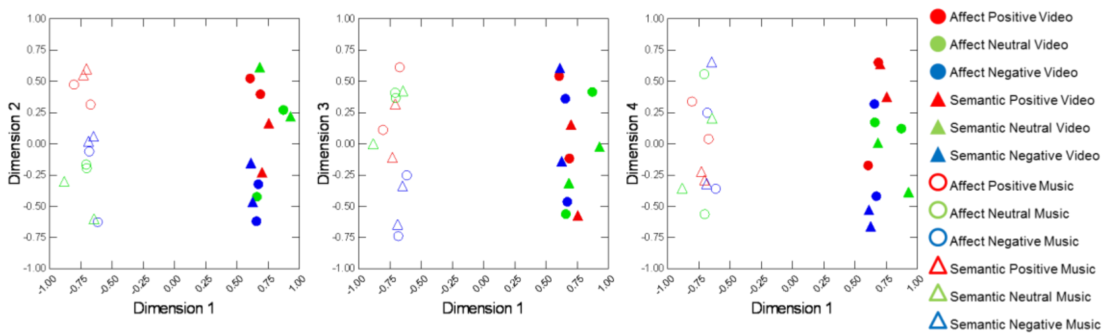

Figure 5.9 Lower dimensional representation of valence based on fMRI data. A four dimensional solution is shown providing modality-general valence information as well as modality-specific valence.

5.3.5. Within-participant classifications based on anatomical ROIs

ROI-based MVPAs were conducted in predefined anatomical ROIs. First, within-modal classifications were performed if anatomical ROIs had valence information for each modality. One sample t-tests revealed that within-modal classification accuracies of all three ROIs were significantly higher than chance for all three types of classifications, *ps* < .001 (Figure 5.10), suggesting that activation patterns for ROIs in frontal areas and the temporal area indeed contain valence information for each modality.

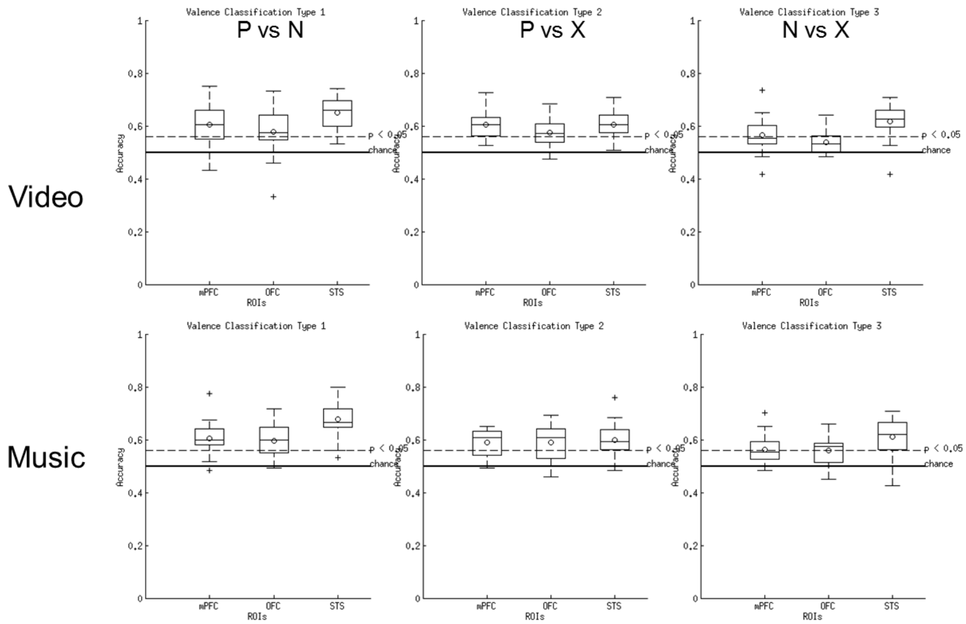

Figure 5.10 Within-participant within-modal classification accuracies based on anatomical ROIs (mPFC, OFC, and STS) for identifying trials as positive vs. negative (left panel), positive vs. neutral (middle) and negative vs. neutral (right), summarized across twenty participants by box plots are shown.

Modality-general processing implies that the valence representation from visual stimuli overlaps with the valence representation from auditory stimuli. Thus classifiers trained from visual stimuli may successfully classify auditory valence and vice versa. ROI-based cross-modal MVPA can be used to confirm modality-general representation of valence in these regions. To examine whether valence can be decoded from predefined ROIs, a logistic regression classifier was trained from video trials and tested on music trials, and vice versa. One sample t-tests revealed that classification accuracies of all three ROIs as well as combined ROIs were significantly higher than chance level for all three types of classifications, *ps* < .001 (Figure 5.11). These results support the assertion that the activation patterns for ROIs in frontal areas and the temporal area indeed represent modality-general valence processing. A repeated measures ANOVA was run to test if there were any classification accuracy differences between the three ROIs.

The result was that cross-modal classification accuracy from the STS was significantly higher than accuracies from the two frontal areas for the two classification types (positive-negative and negative-neutral). This result suggests that even though frontal areas may have representation of modality-general affective states, that representation may be weaker than that in the STS.

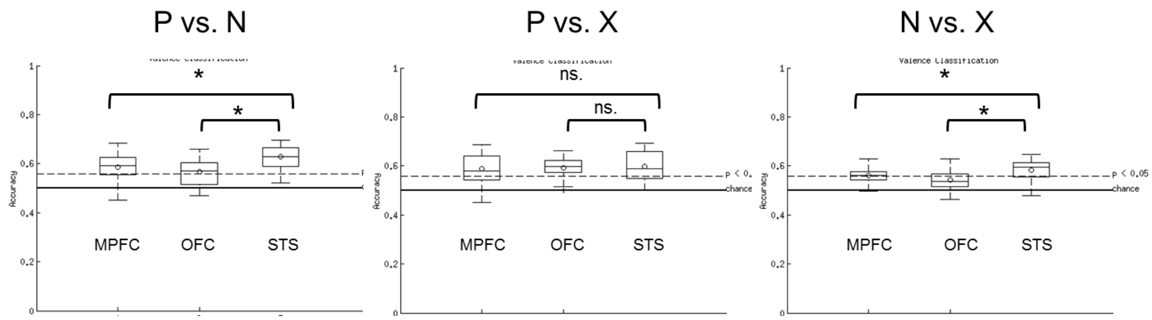

Figure 5.11 Within-participant cross-modal classification accuracies based on anatomical ROIs (mPFC, OFC, and STS) for identifying trials as positive vs. negative (left panel), positive vs. neutral (middle) and negative vs. neutral (right), summarized across twenty participants by box plots are shown.

5.3.6. Lower dimensional representation of affective space based on anatomical ROIs

To examine the lower dimensional representation of affective space within each ROI, the correlation matrices for odd and even trials of each of the experimental conditions were generated for each ROI and each participant. The general procedure was the same as the MDS from whole brain but the only difference was that all of gray matter voxels in each ROI were used. Figure 5.12 shows the MDS result for each ROI. The representation from mPFC showed the first dimension reflecting modality-general valence separating positive, neutral, and negative stimuli from both modalities, the second dimension reflecting music valence separating valence categories only from music trials, and the third dimension reflecting video valence separating valence categories only from video trials. There was no information regarding modality from mPFC. Similarly, the MDS result from OFC showed no modality information but revealed dimensions for music valence, video valence, and modality-general valence, respectively. The MDS result from STS showed strong information of modality first, separating visual and auditory trials. The second through fourth dimensions reflected modality-general and modality-specific valence. Overall, the MDS results provide a clear visualization of the data, with modality, modality-general valence as well as modality-specific valence aspects of the stimuli separated in the lower dimensional space and related in expected ways to the neural responses at the whole brain level. The lower dimensional solutions provide additional support for the result of within-participant cross-modal classification based on ROIs, which suggested modality-general representation of affective states.

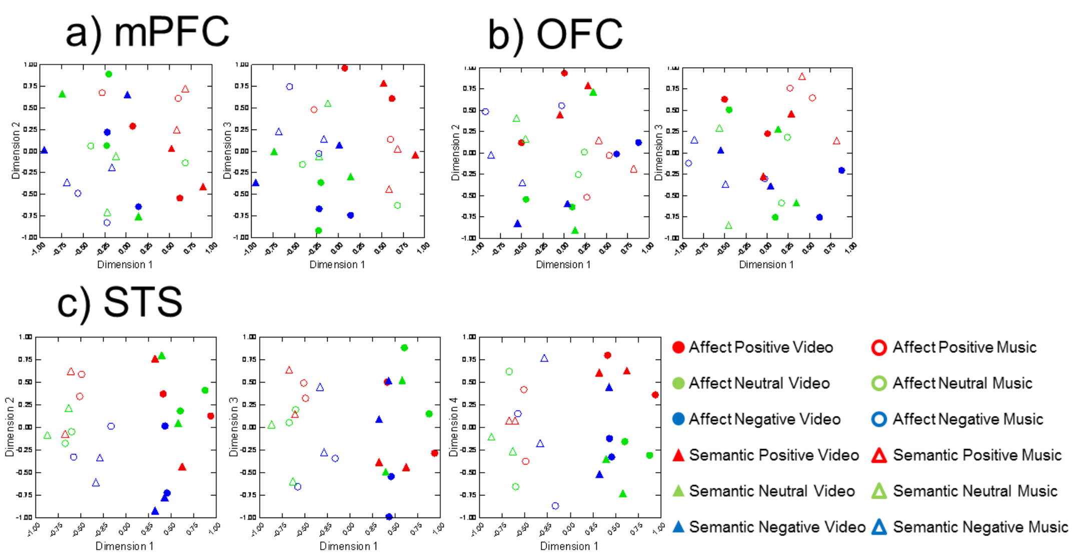

Figure 5.12 Lower dimensional representations of valence based on fMRI data for a) mPFC, b) OFC, and c) STS. Both frontal areas do not show modality information but have modality-general valence information as well as modality-specific valence. MDS results from STS show modality-general and modality-specific valence information as well as modality.

5.3.7. Cross-participant classification

To examine the consistency of the neural representations of affect across participants, the whole brain activation data from all but one participant were used to identify the valence category of stimuli presented to the excluded participant. This test was performed for three two-way classifications as described in within-participant classification section. Unless otherwise stated, classifications were based on 600 the most replicable voxels across the participants in the training set[3].

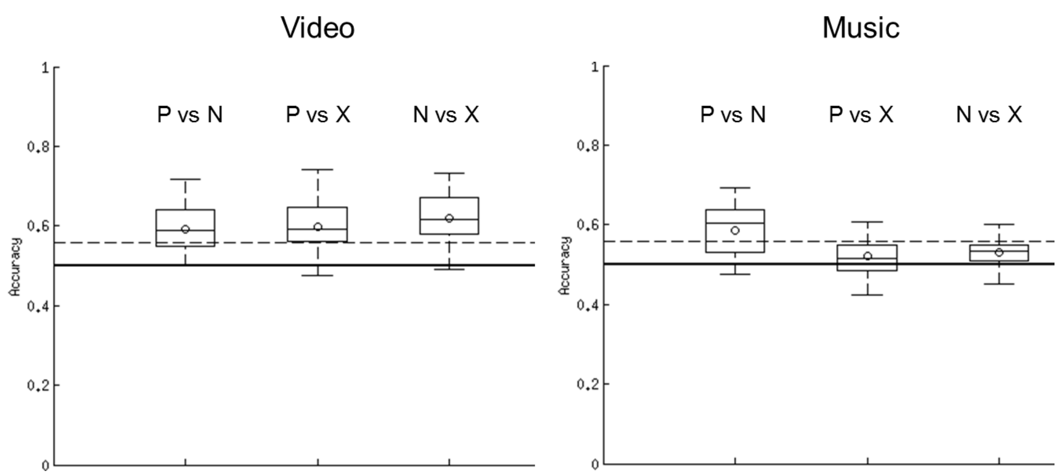

Figure 5.13 Cross-participant within-modal classification accuracies based on whole brain for video (left panel) and music (right panel) trials. Three classifications are shown as positive vs. negative (left box plot), positive vs. neutral (middle), and negative vs. neutral (right), summarized across twenty participants by box plots.

First, cross-participant within-modal classifications were performed to demonstrate if there is commonality of modality-specific valence representation across participants. The results showed that the mean accuracies for three types of the classifications of video trials were .59, .60, and .62, and three one-sample t-tests revealed

---

[3] In order to examine the difference between two numbers of voxels included in the cross-participants MVPA, within-modality (video and music, separately) classification accuracies based on 600 and 3000 voxels were compared. The results were that in both modalities, the classification accuracies based on the two voxel numbers were not significantly different, *ps* > .05. For the computational convenience, 600 voxels were used for the further analyses.

that these three sets of the accuracies were significantly higher than the chance level, $p < .001$ (Figure 5.13). Based on binomial distribution, the number of significant participants was 15, 16, and 16 out of 20, respectively. Similar but weaker results were found for the music trials. The mean accuracies for three types of the classifications of music trials were .59, .52, and .53 and three one-sample t-tests revealed that positive vs. negative, and negative vs, neutral accuracies were significantly higher than chance, $p < .05$, but classifying positive vs. neutral for the music trials was not significant, $p > .05$. Based on binomial distribution, the number of participants with significant classification accuracies was 13, 4, and 4 out of 20 for these three comparisons, respectively.

Next, cross-participant combined-modality MVPA was run to test the commonality across participants for valence information ignoring the modality distinction. When classifiers were trained to identify valence categories from both modality trials, the mean accuracies across participants, .56, .54, and .56, were significantly greater than chance for all three types of classifications, $p < .001$ (Figure 5.14). Based on binomial distribution, the number of participants with classification accuracies significantly above chance was 12, 9, and 11 out of 20, respectively. This result demonstrates that classifiers trained on other participants' data were able to identify valence reliably above chance for the excluded participant. Thus, the neural activation patterns elicited by affective categories have some consistency across individuals, implying that the neural representation of affect is similar from one individual to another.

Finally, cross-participant cross-modality classifications were performed to determine if the modality-general representation of valence is common across participants. Specifically, classifiers were trained on all video trials of 19 participants and tested on music trials of the excluded participant. The same procedure was repeated for training on music trials and testing on video trials. The cross-participant cross-modality classification, however, was not successful for any type of classification, *ps* > .05. This result does not support the conclusion that modality-general representation of valence is consistent across participants.

Overall, three sets of cross-participant MVPA results revealed that visual-specific, auditory-specific, and combined-modality valence representations were consistent across participants, whereas this was not found for the representation of modality-general valence. This result does not mean that there is a lack of modality-general representation of valence because within-participants cross-modal classifications at the whole brain level as well as anatomical ROIs were successful. The two sets of results (successful within-participant cross-modal classifications and *un*successful cross-participant cross-modal classifications) may imply that modality-general representation of valence may be more varied across individuals than modality-specific representation of valence.

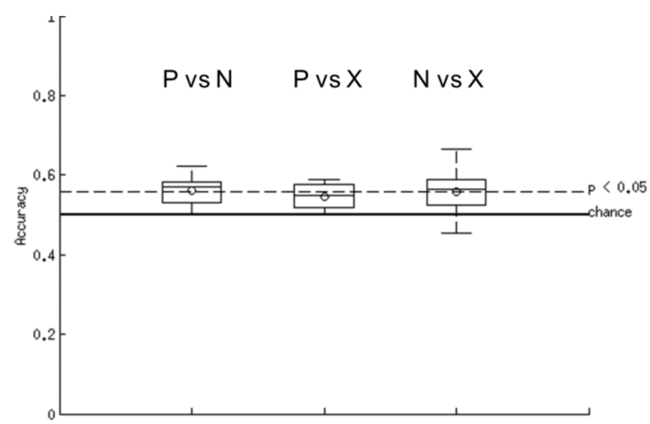

Figure 5.14 Within-participant combined-modal classification accuracies for identifying trials as positive vs. negative (left), positive vs. neutral (middle) and negative vs. neutral (right), summarized across twenty participants by box plots are shown.

5.3.8. Interim summary

These results offer support to the premise that general valence information is represented at the whole brain level as well as at predefined anatomical ROIs. The within-participant decoding results demonstrate that information unique to valence lies within distributed patterns of brain activation across the whole brain and can be used to predict which valence levels a participant was experiencing as elicited by viewing of silent movies or musical clips. Within-participant classification was above chance for the majority of the participants. Cross-modal classification revealed that there is modality-general representation of valence from activity patterns of the whole brain, frontal areas, and a temporal region. MDS results support the results from MVPA that showed modality-general representation of valence. The commonality of modality-specific valence representation was confirmed by cross-participant classification.

5.3.9. Searchlight analysis

The two separate searchlight analyses were performed to pinpoint the clusters involved in 'modality-general processing of valence' and 'modality'.

Valence based similarity pattern between modalities

Table 5.2. Significant clusters from a searchlight analysis ($p < .01$, FWE corrected, cluster size > 50).

			MNI coordinates				
Anatomical region	Hemisphere	Cluster size	x	y	z	T	Z
Transverse Temporal gyrus	R	94	51	-10	6	11.31	6.17
Superior Temporal gyrus	L	97	-51	-23	16	9.79	5.78
Middle Temporal gyrus	R	82	51	-72	6	8.86	5.51

First, searchlight analysis was performed to find significant valence based similarity patterns between the two modalities. The correlation between the two modalities was computed for all voxels in a 5×5×5 cube, and individual z-transformed maps were submitted to a random-effects group analysis. The analysis revealed three clusters: the right transverse temporal gyrus, the left superior temporal gyrus, and the right middle temporal gyrus ($p < .01$, FWE corrected, cluster size > 50; Figure 5.15, Table 5.2). The superior temporal gyrus and middle temporal gyrus have been previously implicated in valence processing (Baucom et al., 2012; Linquist et al., 2015; Peelen et al., 2010), found by multivoxel pattern analysis. The transverse temporal gyrus is also involved in emotional experience (Habel et al., 2005) or affective states evoked by music (Koelsch et al., 2006; Omar et al., 2011).

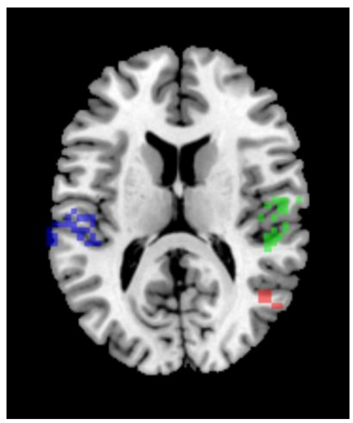

Figure 5.15 Results of the whole brain searchlight analysis. Three brain regions showed modality-general representation of valence. Axial slice of the brain is showing three clusters (right transverse temporal gyrus: green, left superior temporal gyrus: blue, and right middle temporal gyrus: red).

Two additional tests were performed to see if those clusters are truly informative of modality-general representation of affect (Etzel et al., 2013). First, cluster-based MVPA was conducted in order to validate the clusters identified by the searchlight analysis. Note that the purpose of the searchlight analysis was to locate the voxels where valence is represented in a modality-general way. It was assumed that if modality-general processing of affect occurs, then the valence representation from visual stimuli may be similar to the valence representation from auditory stimuli. Thus it is possible to hypothesize that the classifiers trained from visual stimuli may be successful to classify the auditory trials and vice versa. Cluster-based cross-modal MVPA confirmed modality-general representation of valence. To examine whether valence can be decoded from the identified clusters, a logistic regression classifier from video trials was trained and tested on music trials and vice versa. One sample t-tests revealed that three classification accuracies (positive vs. negative, positive vs. neutral, and negative vs. neutral) of all three clusters as well as the combined cluster were significantly higher than chance, *ps* < .001 (Figure 5.16). Binomial significance tests indicated that the majority of participants were significantly higher than chance for positive versus negative conditions for each cluster (16, 17, and 14 out of 20, respectively). These results suggest that modality-general valence is indeed represented in the clusters identified by the searchlight analysis.

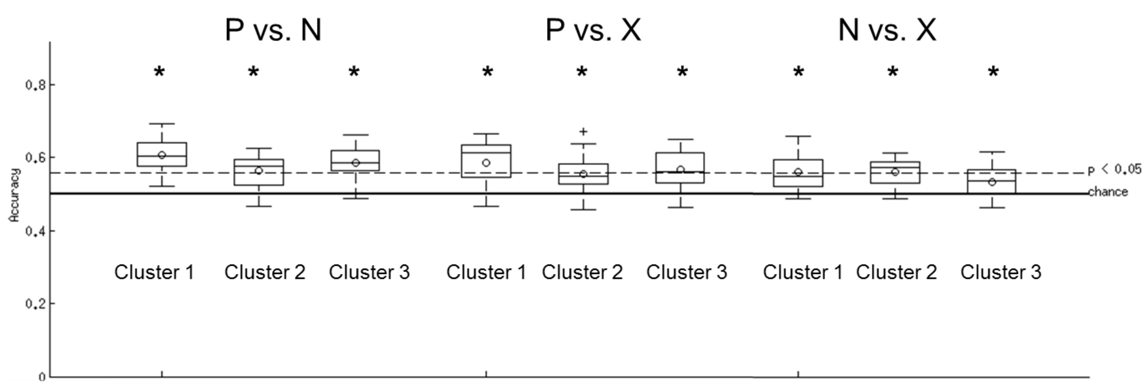

Figure 5.16 Cluster-based cross-modal classification accuracies based on searchlight analysis (Cluster 1: the right transverse temporal gyrus, Cluster 2: the left superior temporal gyrus, and Cluster 3: the right middle temporal gyrus) for identifying trials as positive vs. negative (left three box plots), positive vs. neutral (middle three), and negative vs. neutral (right three), summarized across twenty participants by box plots are shown. * $p < .05$

The second validation was conducted using MDS. In order to illustrate the representation of valence in the clusters identified by searchlight analysis, the similarities between voxel response patterns were evoked by the 60 exemplars using an MDS. The four dimensional models described the representations for each cluster as well as the combined cluster (Figure 5.17). Overall, the four dimensional representations were consistent across all three areas. Modality dimension emerged first, followed by modality-general valence and modality-specific dimensions. The key finding is that dimension 2 results for all three areas confirmed that the clusters identified by searchlight analysis are informative of modality-general valence representation.

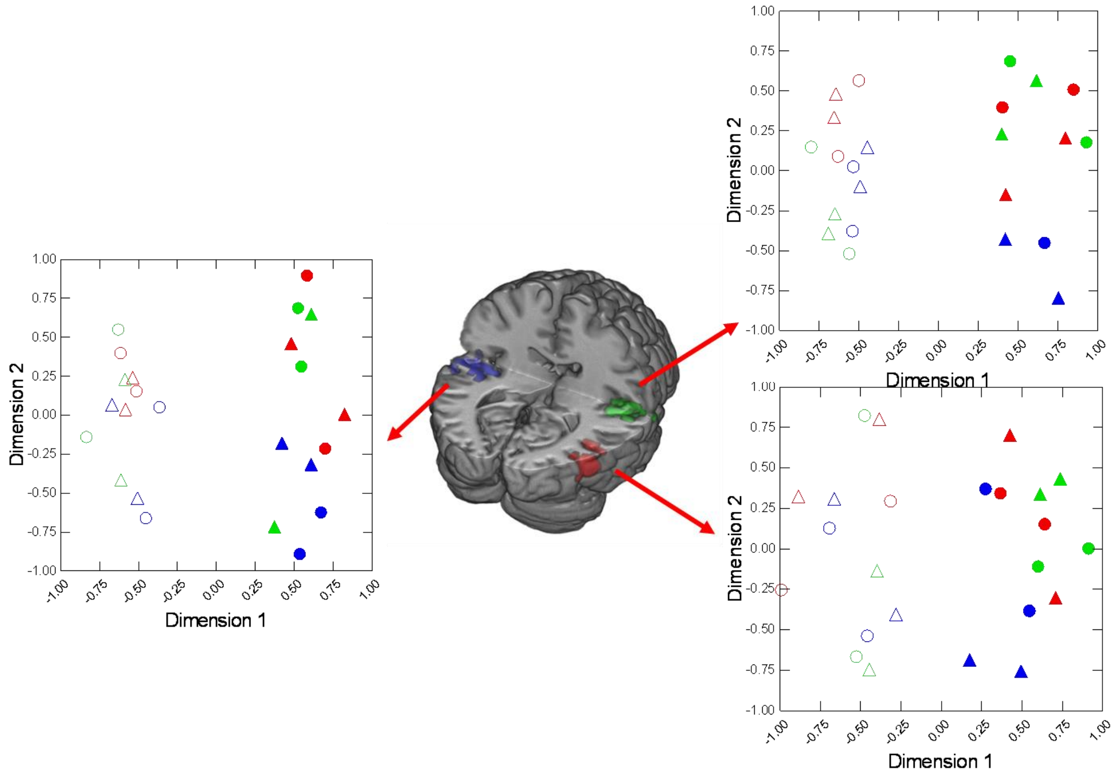

Figure 5.17 Cluster-based MDS results. Multidimensional scaling solutions from three clusters identified by searchlight analysis are showing modality and modality-general valence information. Only dimension 1 versus dimension 2 plots are shown.

Identifying modality-specific regions

The purpose of the second searchlight analysis was to identify brain regions involved in modality information. It was hypothesized that modality-congruent valences will be successfully decoded. However, it was also investigated whether modality-incongruent valences would be identified, supporting the idea of cross-modal encoding. The result of the second analysis will tell us if early sensory region activity reflects sensory stimulation *per se* or rather perceptual ‘experience’, since this latter classification does not use any of the sensory features in that brain region.

Figure 5.18 shows three clusters identified from the whole brain searchlight analysis ($p < .05$, FWE corrected, cluster size > 50). The visual specific clusters (red) were located in the occipital lobe, the auditory specific clusters (blue) were located in the bilateral temporal lobe, and between-modal clusters (green) were found in the precentral and middle cingulate cortex.

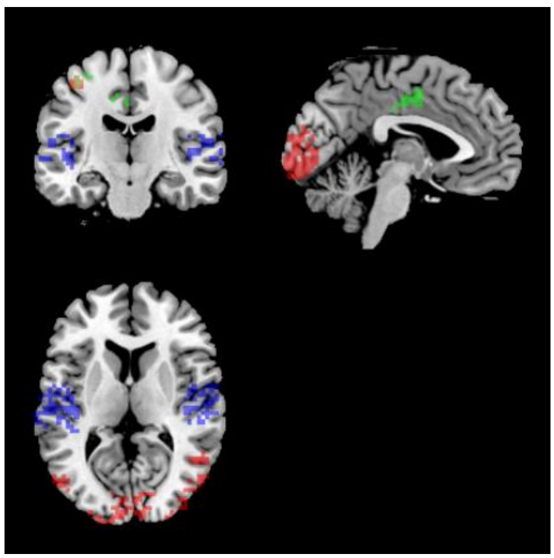

Figure 5.18 Whole brain searchlight analysis identification of modality-specific clusters. Three clusters (visual-specific: red, auditory-specific: blue, and between-modal: green) are shown on coronal (top left), sagittal (top right), and axial (bottom) slices.

Follow up within-modality MVPA was conducted for clusters containing valence information for that modality. For within-modality classification, classifiers were trained and tested on the same modality trials. Cluster-based MVPA confirmed modality-specific representation of valence (Figure 5.19). Not surprisingly, one sample t-tests revealed that for three classifications (positive vs. negative, positive vs. neutral, and negative vs. neutral), visual valence was successfully decoded from the visual-specific region and auditory valence was successfully decoded from the auditory-specific region, *ps* < .001. Binomial significance testing for each individual indicated that the majority of participants were significantly more accurate than chance (19, 20, 20 for visual trials and 18, 17, 17 for auditory trials). These results suggest that the two modality-specific regions are informative to modality-specific valence.

The second research question was if valence information is coded in the modality-incongruent region: visual valence information in the auditory-specific region and auditory valence information in the visual-specific region. The within-modality MVPA result indicated that for all three classifications types, visual valence was successfully decoded from the auditory-specific region and auditory valence was successfully decoded from the visual-specific region, *ps* < .001. In each modality-specific region, the accuracy of modality-congruent trials was significantly higher than modality-incongruent trials (video > music in visual region and music > video in auditory region), suggesting that valence information of modality-congruent trials is stronger than that of modality-incongruent trials. The successful within-modal classifications from the modality-incongruent regions suggest that activity even at the early stages of sensory processing could represent inconsistent modality valence as well as consistent modality. Another possibility is that there is modality-general valence processing in both regions. To test this possibility, cross-modal classification was performed on visual- and auditory-specific regions. However, a classifier trained from one modality was not able to decode the other modality, *ps* > .05. Therefore, although valence for both modalities is coded in both sensory related regions, these regions do not code valence in a modality-general way.

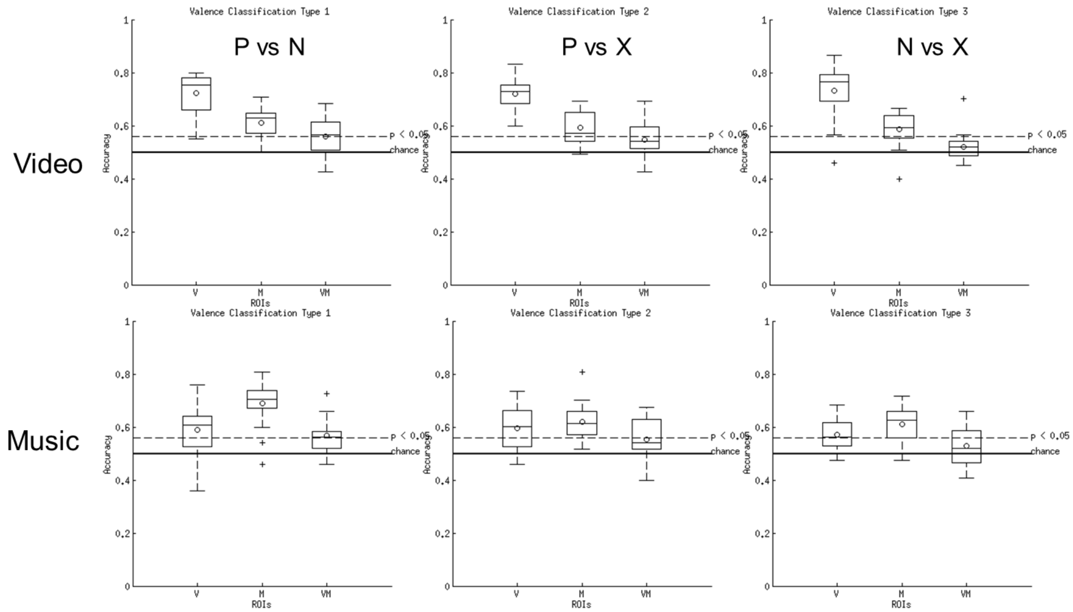

Figure 5.19 Cluster-based within-modal classification accuracies based on searchlight analysis (V: visual-specific cluster, M: music-specific cluster, VM: between-modal cluster), summarized across twenty participants by box plots. Classifications were to identify trials as positive vs. negative (left), positive vs. neutral (middle) and negative vs. neutral (right), trained from and tested on video trials (top) and music trials (bottom).

One possible critique to the result of this analysis is that these modality-specific regions may overlap with three modality-general clusters identified by the first searchlight analysis. Because those three clusters were also found in the temporal and inferior part of the occipital region, it can be argued that the 'modality-general' voxels in those areas might work for the within-modal classification for the opposite modality trials. In order to rule out this possibility, anatomically defined ROIs known to be highly modality-specific were tested with the same procedure. Specifically, masks of primary visual cortex (V1) and primary auditory cortex (A1) were created from the SPM Anatomy Toolbox. The results are shown in Figure 5.20. Again, visual valence was successfully decoded from V1 and auditory valence was decoded from A1, supporting the results of the previous analysis. The within-modality MVPA result indicated that for all three classifications types, visual valence was successfully decoded from the primary auditory cortex and auditory valence was successfully decoded from the primary visual cortex, *ps* < .001. Again, classification accuracies of visual valence from the primary visual cortex and auditory valence from the primary auditory cortex were significantly higher than those of modality-incongruent trials (video > music in visual region and music > video in auditory region), though the inconsistent trials were significantly higher than the chance level. Again, cross-modal classifications were conducted to test if there is modality-general processing of valence from both regions. The results showed that classification accuracies from both regions were not successful, *ps* > .05, suggesting that these regions do not encode valence in a modality-general way.

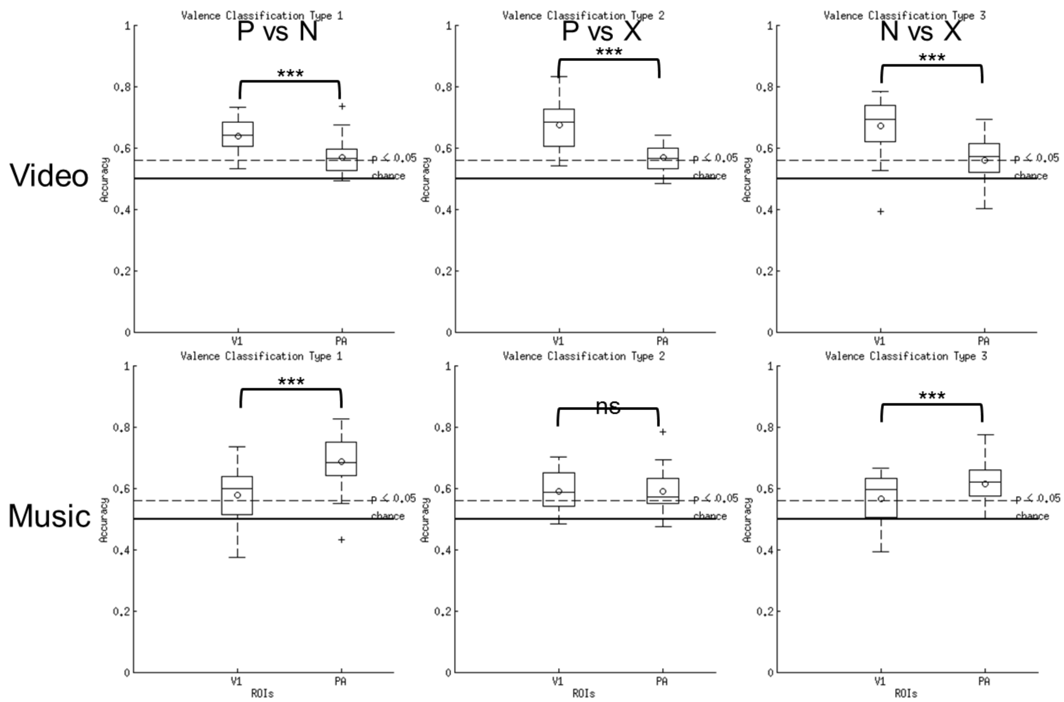

Figure 5.20 ROI-based within-modal classification accuracies (V1: primary visual cortex, PA: primary auditory cortex), summarized across twenty participants by box plots. Classifications were to identify trials as positive vs. negative (left), positive vs. neutral (middle) and negative vs. neutral (right), trained from and tested on video trials (top) and music trials (bottom).

5.3.10. Representational similarity analysis

An RSA was conducted to test which specific conceptual model combinations can explain fMRI data from the whole brain, anatomical ROIs, and three clusters identified by the searchlight analysis. For the whole brain, the final model included four parameters modality, vv2a, vv3c, and mv2a (R^2 = .664). The strongest parameter was modality, followed by visual valences and auditory valence. The pattern that the stronger estimate of visual parameters over auditory parameters was consistent with the result of within-modality MVPA that showed higher classification accuracies from video trials than music trials. However, no modality-general parameters were significant, even though cross-modal classification was successful at the whole brain level.

For the mPFC, the final model included one parameter; mgv3c (modality-general valence) (R^2 = .006). The significant modality-general valence parameter is consistent with the successful cross-modal classification result. The model testing for OFC was similar to mPFC. The significant parameter for OFC was mgv2b (modality-general valence) (R^2 = .004). None of the modality-specific valence models were significant for the mPFC and OFC.

For the STS, the final model included four parameters: modality, vv1 (bipolar visual valence), vv3b, and vv3c (general visual valence) (R^2 = .363). The strongest parameter was modality, followed by bipolar model and general valence models. None of the modality-general parameters were significant.

The nonlinear models for the bilateral temporal clusters (the right transverse temporal gyrus and the left superior temporal gyrus) identified by the searchlight analysis were very similar to that of the STS. Both of the models for the temporal searchlight clusters included four parameters: modality, bipolar visual valence, and general visual valence (positive over negative/negative over neutral and negative over positive/positive over neutral) (R^2 = .551, .488). The final model of the other cluster (the right middle/inferior temporal gyrus) included four parameters: modality, bipolar auditory valence, and general auditory valence (positive over negative/negative over neutral and negative over positive/positive over neutral) (R^2 = .279). Interestingly, both groups of STS/two temporal searchlight clusters and one middle temporal searchlight cluster supported both bipolar and general valence hypotheses of valence for visual and auditory stimuli, respectively.

The results of RSA for the mPFC and OFC were consistent with the previously described analyses. Both the cross-modal classification and MDS results suggested that there is modality-general valence processing of valence from mPFC and OFC, which was also confirmed by significant parameters of modality-general models in RSA. The modality information was not found in MDS and RSA as well. However, the results of RSA for the whole brain and the STS were not entirely consistent with the MVPA and MDS results. Both MVPA and MDS support modality-general processing of valence evidenced by a successful cross-modal classification and a common valence dimension of visual and auditory trials. However, the final models of the whole brain, STS, and three searchlight clusters did not include any modality-general valence models. One possible explanation is that the RSA may not have enough power to test valence hypotheses. Though the final models of the whole brain, STS, and three searchlight clusters are fairly high, the variances explained by the model were heavily driven by ‘modality’. For example, the reduced model for the whole brain with three parameters without modality had $R^2 = 0.290$ so that the variance that is explained by modality given the other three parameters is .374. Please note that the overall R^2 of the final models of the mPFC and OFC were extremely low ($R^2 = .006, .004$) though in both regions the modality-general valence models were significant parameters. In sum, the RSA supported the modality-general processing of valence for the mPFC and OFC (Lindquist et al., 2015) as shown by cross-modal MVPA and MDS. It also suggests that it may be necessary to consider modality information when testing a valence hypothesis.

5.3.11. Effect of task

We hypothesized that modality-general affective processing may depend on the attentional focus as manipulated by a judgment task. Specifically, it was hypothesized that intentional evaluation of the affective aspect of the stimuli will result in modality-general representation of valence, whereas focusing on semantic features of affective stimuli will result in modality-specific representation of valence only. Representations of valence under two tasks are compared by cross-modal MVPA and MDS.

First, within-modal classification was run for each task (Figure 5.21). For the affective task, within-modal classifications accuracies of both video and music trials were significantly higher than the chance level, *ps* < .001, with the exception of the affective tasks-music trials-positive vs. neutral classification. For the semantic task, all of within-modal classifications of both of video and music trials were significantly higher than the chance level, *ps* < .001. These results suggest that under each task, within-modal valence information can be predicted. Next, the repeated ANOVA revealed no significant difference between tasks, $F(1,19) = .06$ $p > .05$; $F(1,19) = 3.03$, $p > .09$; $F(1,19) = .38$, $p > .05$. The main effect of modality was significant for the positive vs. neutral, $F(1,19) =$

10.02, $p < .01$, and negative vs. neutral classification, $F(1,19) = 13.33$, $p < .01$. The two modalities were not significantly different for positive vs. negative classification, $F(1,19) = 2.92$, $p > .05$. The interactions between modality and task were not significant, *ps* > .05. In sum, the within-modal MVPA results suggest that there is no effect of task on valence classification within-modalities.

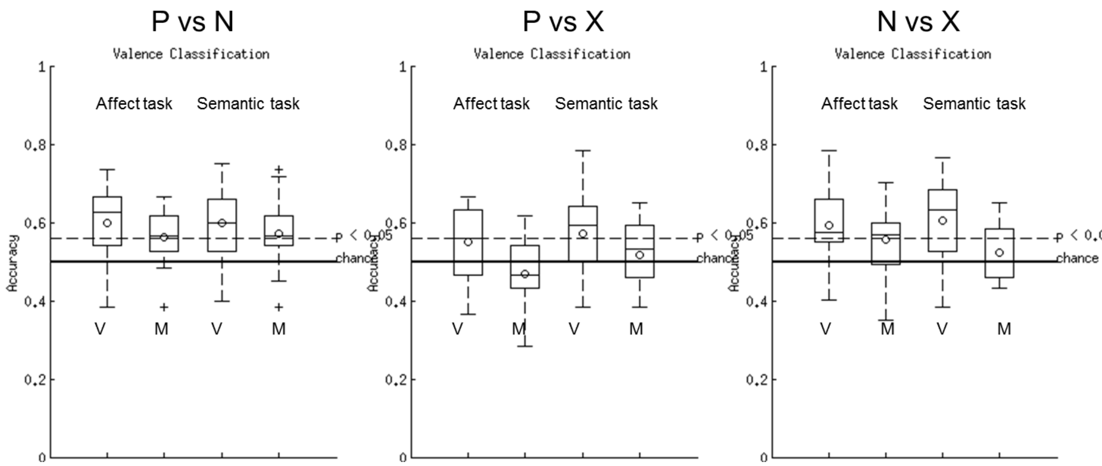

Figure 5.21 Within-modal classification accuracies for each task. Modalities were denoted by V (video) and M (music), summarized across twenty participants by box plots. Classifications were to identify trials as positive vs. negative (left), positive vs. neutral (middle) and negative vs. neutral (right).

Note that in the behavioral data for fMRI participants, response accuracies decreased from positive to negative to neutral conditions when the modality was auditory, but that the negative condition was more accurate than neutral and positive conditions when the modality was visual. This finding suggests that valence differences of classification accuracy could be confounded with differences in task difficulty. This possibility was tested by comparing within-participants within-modality classifications of three comparisons (positive vs. negative, positive vs. neutral, and negative vs. neutral) with two assumptions: 1) for video trials, positive vs. negative and negative vs. neutral classifications should be more accurate than positive vs. neutral, and 2) for music trials, positive vs. neutral classification should be more accurate than negative vs. neutral and positive vs. negative classifications. The repeated-measures ANOVA revealed that the three accuracies were not significantly different for rach modality, *ps* > .05, suggesting that the effect of the behavioral accuracy difference between three valence categories of both modalities on the fMRI MVPA was minimal.

The cross-modal classification accuracies under two tasks at the whole brain level were compared and the result showed that there was no difference between the two tasks for all three types of the classifications, *p* > .05 (Figure 5.22), though under both tasks, the cross-modal classifications were significantly higher than the chance level, *ps* < .001. Both of the multidimensional scaling solutions under two tasks at the whole brain level showed modality-general dimensions. Finally, the same cross-modal classifications were performed at three clusters identified by the first searchlight analysis, and the results also revealed no difference between the two tasks. In sum, within- and cross-modal MVPA and MDS results suggest that there is no difference in representations of valence under the two types of tasks, with data from both tasks supporting significant valence classification within and across-modalities.

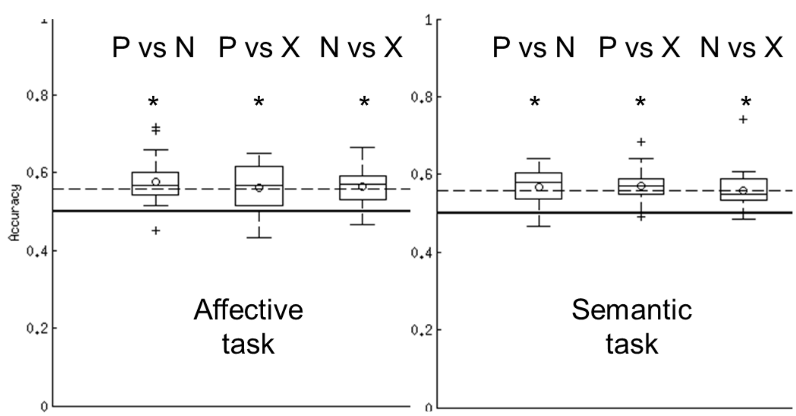

Figure 5.22 Cross-modal classification accuracies for affective (left panel) and semantic (right panel) tasks. Classifications were to identify trials as positive vs. negative (left), positive vs. neutral (middle) and negative vs. neutral (right) for each panel.

5.3.12. Supplementary analyses

Difference of two modalities

In the current study, within-modality classifications were performed if there was valence information at the whole brain level or multiple ROIs. One interesting research question is if there is relative dominance between the two modalities. Specifically, the dominant role of vision over audition for emotion processing has been proposed (Klasen et al., 2014). One can point out that these stimuli differed not just in modality (visual / auditory) but in episodic structure (visual stimuli could be interpreted according to a meaningful sequence of events, whereas the auditory stimuli in the current stimuli set were music, which is non-episodic). This difference in episodic structure may lead to a difference in classification accuracy. For example, one can argue that video trials can be decoded with greater accuracy than musical trials because video trials are stronger and more meaningful. This can be tested with the within-modality classifications. A series of repeated-measures ANOVA found significantly higher accuracy from video stimuli over music stimuli for the whole brain (video: $M = .75$, music: $M = .66$), and the left superior temporal gyrus (video: $M = .62$, music $M = .54$), and significantly higher accuracy from music stimuli over video stimuli for the right transverse temporal gyrus (video: $M = .57$, music: $M = .60$), and the right middle temporal gyrus (video: $M = .55$, music $M = .59$). Please note that the results of RSA indicated these regions' final model included significant modality parameters. Repeated measures ANOVAs from mPFC and OFC revealed no significant effect of modality on classification accuracy, *ps* > .05. Again, the final models of these two frontal regions did not have modality parameters, which means that the brain regions that do not have modality information also do not have modality valence difference. When repeated measures ANOVAs were run on all 7 regions at the same time, the main effect of modality was not significant, $F(1,19) = 1.12$, $p > .05$. Though it has been reported that visual information is stronger than auditory in memory (Cohen et al., 2009), perception (Kumpik et al., 2014), and facial reactions (Sestito et al., 2013), these classification results suggest that modality dominance may be dependent on the specific brain regions.

Classification types

Because there are three categories in valence, three different two-way classifications were performed for all of the MVPAs in the current study. One interesting research question is if there are any differences between the three classifications. Again, a repeated measures ANOVA was run on the within-modality classification accuracies. The results were that the main effect of classification types was significant, $F(2,38) = 16,79$, $p < .001$, and interaction between classification type and ROI was also significant, $F(2,38) = 2.52$, $p < .01$. A set of separate ANOVAs revealed significant main effects of classification type from each of the region, *ps* < .001. When comparing means of the accuracies, the highest classification accuracy was from positive vs. negative ($M = .628$), followed by positive vs. neutral ($M = .604$), and negative vs. neutral ($M = .589$). Given the fact that most of the features were matched across valence categories (low level features, arousal, and number of trials), the significant differences between valence categories should be mainly attributed to valence. These results suggest that the difference between positive and negative is the biggest out of the three comparisons. One possibility is that neutral trials can be misclassified to either positive or negative categories, whereas positive or negative categories are hard to be misclassified to the neutral category. This argument can be supported by the mean standard deviations of valence ratings from fMRI participants for each valence category (negative *SD*: 1.495, neutral *SD*: 1.70, positive *SD*: 1.39), showing that positive and negative stimuli were less variable, whereas neutral ones were more variable. One way to look at this more carefully will be to examine three-way classifications and see how each valence category is *mis*classified.

Psychopathological studies have shown that depressed people selectively attend to the negative-valence stimuli and memories because depression has been found to be associated with difficulties inhibiting the processing of negative material (Joormann & Gotlib, 2007). This was supported by neuroimaging studies (i.e. Surguladze et al., 2005). Thus, it is expected that depressed people may have a better performance for positive vs. negative and negative vs. neutral classifications compared to positive vs. neutral classification because negative stimuli may be more easily distinguishable from the other types of valence. Note that though Habes et al. (2013) compared three types of classifications of depressed people using IAPS pictures and reported fairly similar accuracies between the three classifications (positive vs. negative: 92%, negative vs. neutral: 86%, neutral vs. positive: 89%), it is not clear that valence manipulation lacks confound of arousal because positive and negative stimuli in IAPS tends to be more arousing than neutral stimuli.

5.4. Summary

The current fMRI study used MVPA methods to identify the valence representations elicited by viewing silent videos and listening to music. These analyses were carried out on the distributed representations over the whole brain elicited by affective stimuli, theoretically defined a priori anatomical ROIs, and clusters revealed by a searchlight analysis. General valence was successfully decoded from patterns of brain activation within participants. The within-participant decoding results demonstrate that information unique to valence lies within distributed patterns of brain activation across the whole brain as well as frontal and temporal regions of interest and can be used to predict which valence levels a participant was experiencing as elicited by exposure to affect-related music or video. Within-participant classification was above chance for the majority of the participants. Also, successful cross participant classification demonstrated that there is a commonality between individuals for valence representation. However, this type of classification does not address the issue of modality-specific and modality-general affect representations. Modality-general processing of valence was tested by cross-modal classification and MDS and both analyses confirmed that valence is represented in a similar way between the two modalities at the whole brain level as well as frontal and temporal regions. Searchlight analysis was conducted to explore the brain for areas of modality-general representation of valence and three clusters including right transverse temporal gyrus, left superior temporal gyrus, and right middle/inferior temporal gyrus were identified. Additional cross-modal classification and multidimensional scaling analyses validated modality-general representation of valence in those clusters. Modality-specific areas were located by additional searchlight analyses: the occipital region for visual stimuli and the temporal region for auditory stimuli were identified. Within-modality classification confirmed that those modality-specific areas are involved in valence processing of the corresponding modality. Interestingly, visual valence was also decoded in the auditory region and auditory valence was decoded in the visual region, though those classifications were less accurate than the modality-congruent accuracy.

This was supported by within-modal classifications in anatomical sensory cortices. This result suggest that the primary sensory regions may encode perceptual experience rather than (or as well as) sensory stimulations. Another part of the hypotheses of the current study was to compare two types of attentional focuses – affective and semantic. It was hypothesized that focusing on the affective aspect of the stimuli may lead to modality-general processing whereas focusing on the semantic aspect may fail to represent valence in a modality-general way. The result revealed that there was no difference between the two tasks.

CHAPTER 6

GENERAL DISCUSSION

6.1. Summary and implications

This study was designed to investigate how valence information generated from different modalities is represented in the brain. There has been evidence for both modality-general processing (Chikazoe et al., 2014; Klasen et al., 2011; Peelen et al., 2010) and modality-specific processing (Shinkareva et al., 2014). The current study attempted to better understand where and how these two types of encoding occurred. To do so, I applied multiple multivariate techniques to the fMRI data from the whole brain, a priori anatomical regions, and clusters found by searchlight analyses. A series of behavioral studies was first run to develop an experimental stimulus set that met requirements of valence stimulus types while holding arousal constant across valence categories. Sixty unique exemplars with two replicates were chosen for the main fMRI experiment. The valence manipulation was successful, with valence categories were equated on arousal ratings. Many previous studies of emotion have failed to unconfound valence and arousal, which is particularly important in assessing comparisons to neutral stimuli that tend to have lower arousal than positive and negative stimuli. A second advantage of the developed stimuli set was that the many low level features of the stimuli were also equated across valence categories, increasing the likelihood that decoding was primarily based on valence information and not some correlated perceptual dimension. Previous behavioral and psychophysiological studies including galvanic skin response have revealed a systematic relationship between visual, motion, and acoustical features with behavioral and physiological responses to affective states (Gabrielsson & Lindstrom, 2001; Juslin & Laukka, 2004; Lakens et al., 2013). The stimulus sets in the current study matched many of the low level features between valence categories so that any difference between experimental conditions can most likely be attributed to the valence of the stimuli and not to the arousal levels or low level features of the stimuli.

The current fMRI study used multiple multivariate analysis tools to analyze the neuroimaging data. General valence was successfully decoded from patterns of whole brain activation pattern within participants. The successful cross-modal classification demonstrated that there is modality-general processing of valence at the whole brain level. Cross-modal classification also provides the strongest argument that valence information rather than correlated perceptual information is being coded, as perceptual based classification would not be expected to be successful in the cross-modal case given the vastly different perceptual features for the two modalities. The MDS results supported these results by showing that there was a common valence dimension for visual and auditory trials as well as visual- and auditory-specific valence dimensions. A successful cross participant classification demonstrated that there is a commonality between individuals for valence representation. The same analyses were applied to the predefined anatomical ROIs (mPFC, OFC, and STS) and revealed modality-general valence processing evidenced by cross-modal classification and MDS. Two searchlight analyses were performed: 1) to pinpoint the regions which show significantly similar patterns between valence representations of videos and music, and 2) to identify the regions that are involved in each modality regardless of valence. The first searchlight identified three significant clusters: right transverse temporal gyrus, left superior temporal gyrus, and right middle temporal gyrus, and validated with cross-modal classification and MDS. The modality-specific regions found by the second searchlight analysis were the occipital region for visual stimuli and the temporal region for auditory stimuli, which is well known. Within-modality classification confirmed that those modality-congruent areas are involved in valence processing of the corresponding modality. Interestingly, each modality's valence was decoded from the modality-incongruent regions. This finding was supported by within-modal classifications in anatomical primary visual and auditory cortices, suggesting that primary sensory regions may encode perceptual experience. These results imply that modality-specific valence valuation, as cross-modal classification did not work in these regions.

The hypotheses also included comparisons between the two types of tasks. However, the results of within- and cross-modal classifications and MDS revealed that there was no difference between the two tasks. Although this was somewhat surprising, it would seem that active processing of the stimulus semantic information results in activation of modality-general processing as well. Previous studies had generally shown greater affective effects with affect related tasks (Cunningham et al., 2004; Hutcherson et al., 2005; Lange et al., 2003; Straube et al., 2004). This difference in degree might be true in this study as well, but there is no basis to think that task focus shifted the type of valence processing.

Modality-general processing of affect

In sum, this fMRI study supported modality-general processing of affect at the whole brain level and predefined ROIs. The clusters identified by searchlight analysis were consistent with the previous studies (left STS: Peelen et al., 2010). Right temporal regions were also known to be engaged in affect processing (Baucom et al., 2012; Lindquist et al., 2015). Modality-general processing of affect in the frontal areas (mPFC and OFC) was confirmed with ROI-based MVPA and MDS in the current study. However, the searchlight analysis in the current study failed to find frontal areas as the modality-general valence representation. This inconsistency may be explained by the results of the ANOVAs on cross-modal classification accuracies from three anatomical ROIs. The ANOVAs revealed significantly higher accuracies from STS compared to those from mPFC and OFC, suggesting that modality-general processing of affect might be stronger in the temporal regions compared to the frontal regions. Thus, the failure of searchlight to find these regions may be due to a smaller effect size for these regions and hence reduced power.

This study supported modality-general processing of valence hypothesis (Chikazoe et al., 2014; Peelen et al., 2010), whereas Shinkareva et al. (2014) failed to find it. The difference may revolve around inclusion of a task versus no task. Although type of task did not influence modality-general processing in the current experiment, it may be that some task is necessary to trigger this type of encoding. Shinkareva et al. (2014) had no task whereas the other studies that found modality-general processing did have a task that required processing the stimulus. Another possibility for this difference revolves around the issue of power. The number of trials of the Shinkareva et al. study was 96, which was smaller than that of the current study (360), the Peelen et al. study (216), and the Chikazoe et al. (228). The small number of trials may not be enough to detect the modality-general processing of affect. Please note that the cross-modal classification is always based on two cross-validations: training on one modality and testing the classifiers on the other modality trials. In the MVPA, the classification accuracy is dependent on the number of trials of the training session, so only 48 trials were used to train classifiers for the cross-modal classification in the Shinkareva et al. study. Another power issue is the number of participants. They also performed searchlight analysis but failed to find significant clusters. However, searchlight analysis includes group analysis using individual maps, so 8 participants may not be enough to show any significant cluster even though searchlight is a powerful tool to locate informative voxels within individuals. One suggestion is to create an individual significance map (statistically thresholded) and report a group map in terms of the proportion of subjects with a significant searchlight at each voxel as Pereira and Botvinick (2011) suggested. Another suggestion is to use permutation testing which is recommended when the sample size is small. Additionally, cross-modal classifications from significant clusters identified by searchlight analysis for each individual map (not from a group map) may reveal successful decoding.

One interesting application of the current results is related to the autistic patients. It has been reported that autistic children are significantly less sensitive to a facial expressions and less reliable across repeated testing (Kennedy & Adolphs, 2012), and the neural substrates of these abnormalities could be the fusiform face area (FFA): the FFA is hypoactive when autistic individuals view facial expressions (Critchley et al., 2000). However, some studies failed to replicate face difficulties (Gepner, Deruelle, & Grynfeltt, 2001). Matsuda and Yamamoto (2015) raised the possibility that this inconsistency might be due to the modality of the stimuli. They compared within-modal and cross-modal “emotion-matching” performances of autistic and typically-developing (TD) children and found no significant difference between the two groups for the within-modal matching task but a significantly better performance of cross-modal matching task of typically-developing children compared to autistic children. They concluded that autistic individuals had difficulty understanding the relationship between affective prosody facial expression stimuli, whereas they had less difficulty understanding the relationships between visual and visual stimuli. These results suggest that modality-specific representation of affective states may be more intact compared to modality-general representation of affect in autistic children. Thus, it is expected that within-participant within-modal classification would be successful, whereas within-participant cross-modal classification would *not* be successful from autistic children’s neuroimaging data. The prediction of successful within-participant within-modal classification is also supported by Kennedy and Adolphs’ (2012) finding that lower dimensional representations of behavioral ratings of facial expressions from autistic and TD children are quite similar. A study using facial expressions, body movement, and vocal stimuli (Philip et al., 2010) found a worse performance of autistic individuals compared to control group for all types of stimuli. Interestingly, they found significant correlations between three types of performances, suggesting that three modality-specific representations of affect may be connected to each other even though modality-general processing is impaired.

Though within-participant within-modal classification is expected to be successful, cross-participant within-modal classification may not be expected to be successful because Matsuda and Yamamoto (2015) reported a higher accuracy variability of autistic children (*SD* = 29.0) compared to that of TD children (*SD* = 3.1) for within-modal matching task, suggesting that even though modality-specific valence processing of affect is relatively intact, how it is represented might be less consistent across individuals. Finally, reliable correlations between face vs. house MVPA classification performance and standardized measures of symptom severity of autism have been reported (Coutanche et al., 2011). It would be interesting to examine how severity depends on different types of affective representation. Coutanche et al. (2011) examined correlations between multiple severity measures with face vs. house classification accuracies but how severity is related to facial valence classification has not been explored. Also, they conducted multiple MVPAs only within fusiform regions, but how facial expressions are represented within frontal (i.e. mPFC and OFC) and temporal (i.e. STS) regions has not been investigated. These questions can be addressed by future work.

Approaches: merits and weaknesses

In this study, multiple multivariate techniques were used to show how valence information is represented in the brain. First, MVPA was performed to test if affective information can be decoded from the various regions of interests (the whole brain, mPFC, OFC, STS, and the clusters identified by the searchlight analyses). Significance testing was based on one sample t-test for the group level and binomial distribution for the individual level. The three types of MVPA were performed, within-modal classification, general valence classification (regardless of modality), and cross-modal classification. Within-modal classification was used to demonstrate that there is clear valence related information in each stimulus set. However, this does not indicate if this is modality-specific or general. If valance classification ignoring the modality fails when the within-modal classification succeeds, then it is likely that valence specific processing is occurring. If this classification is successful, it can be due to modality-specific and/or modality-general. One of the weaknesses of the within-modal classification and general valence classification is that low-level features may confound the valence classification results. Though many of the low-level visual, motion, and auditory features of the stimuli in the current study were matched, valence classification may be confounded especially for the within-modal classification. However, cross-modal classification rules out this possibility because there are few converging low-level features between the two modalities. The current study demonstrated successful cross-modal classifications from redefined anatomical regions and searchlight clusters, and this suggests that valence is what is being classified. One of the weaknesses of the MVPA is that this technique may not show the internal structure or representation of valence.

To compensate for this deficit, MDS can be used to investigate the internal structure. With this technique, it was easy to visualize how experimental conditions were presented on a lower dimensional space. The extracted dimensions provided additional information regarding the modality-general or modality-specific valence processing of affect. However, MDS is basically a visualization technique, so it lacks a significance testing. We attempted to compensate for this by conducting significance tests on dimensional values from the MDS analysis. This was successful, though not particularly powerful. A future simulation study should evaluate the statistical validity of this type of testing.

Searchlight analysis is a multivariate technique to pinpoint the activation pattern of interest. It can be either classification- or correlation-based. In the current study, two searchlight analyses were run based on correlations between- and within-modalities. One of the flaws of the searchlight analysis is that it can depend heavily on searchlight size/shape and the location of the highly-informative voxels (Etzel et al., 2013). For example, a smaller size of searchlight may be less powerful to detect the informative voxels whereas a bigger size of searchlight may mark truly less-informative voxels as informative. In the current study, the size and shape of the searchlight was fixed as 5×5×5 voxels in a cube. Etzel and colleagues (2013) have suggested applying additional 'confirmatory' analyses from the clusters identified by the searchlight analysis. In the current study, within- and cross-modal classifications and MDS were performed in the searchlight clusters to validate the results. The other issue of the searchlight is the group analysis. In the current study, individual transformed z maps were submitted to SPM level 2 analysis (random effect group analysis) and familywise error was corrected with cluster sizes greater than 50. For the group analysis, recent studies tend to recruit approximately 20 participants (i.e. 16 for Kim et al., 2015, 18 for the Peelen et al., 2010, 30 for Looser et al., 2012). Alternatively, the individual maps can be statistically thresholded and the group-level map can be reported in terms of the proportion of subjects with a significant searchlight at each voxel (Pereira & Botvinick, 2011; Kassam et al., 2013). A permutation-based significance test has also been proposed (Kriegeskorte et al., 2006) and applied (Kim et al., 2015; Oosterhof et al., 2010; Peelen et al., 2010).

Finally, RSA was applied to test different valence hypotheses. The RSA partially supported findings of Lindquist et al. (2015) that positive and negative valence are supported by a flexible set of valence-general regions. At the anatomical mPFC and OFC regions, modality-general parameters were significant, which was consistent with the

ROI-based MVPA and MDS results. The final models for the STS and the searchlight clusters included both bipolar and general valence hypotheses for each modality. Lindquist et al. (2015) found portions of the ventromedial prefrontal cortex, and ACC may serve as candidate ROIs for the bipolarity hypothesis. But they were only able to test if these two areas increased more as positive affect increased but were unable to test if it also showed decreasing activity during negative affect. The RSA in the current study is based on pairwise distance, so it is also not able to test the direction of the bipolarity. For the future study, how to test different valence hypotheses should be considered in greater depth.

Attentional focusing on affective or semantic aspects of the stimuli

One of the hypotheses of the study was to compare two attentional focuses. Specifically, it was hypothesized that affect focusing may lead to modality-general processing whereas non-affect focusing may lead to modality-specific processing of affect. However, the cross-modal classification results revealed no difference between the two tasks. Rather, under both types of tasks, cross-modal classifications were successful, suggesting that even when participants do pay attention to the semantic aspects of the stimuli, which are irrelevant affective aspects, the affective information was encoded regardless of the modalities.

One may raise the argument that the lack of affective processing differences between two types of tasks might be because participants did not properly focus on the task. However, the overall response accuracy was fairly high (M = .79) and comparable to that from the Behavioral Study participants (M = .78), suggesting that fMRI participants were actively engaged in the task in an appropriate manner.

A possibility to explain the lack of difference between tasks is automaticity of affective information. There has been a debate concerning how automatic affect processing is. On the one hand, there are findings showing the absence of activation in the amygdala and visual areas to threat-related visual stimuli during exhaustion of attentional resources (Bishop et al., 2007; Pessoa et al., 2002; Straube et al., 2007). These findings suggest that emotional information may not be processed automatically. On the other hand, it was argued that emotional information, especially threat, is vitally important for the organism (LeDoux, 1998) and can be processed automatically. For example, amygdala activations were found even when participants' attention was distracted from facial expressions (Vuilleumier et al., 2001) or even when the presentation of the stimuli was below the threshold of conscious perception (Whalen et al., 2004). Eye tracking studies (Calvo & Lang, 2004; Nummenmaa et al., 2006) also found a greater probability to have the first fixation and a longer dwelling time on emotional pictures compared to non-emotional pictures, even when the pictures were presented parafoveally. These studies suggest that emotional information may be processed even when the attention was not directed on the stimuli. The current study experimentally manipulated participants' attentional focus on either the affective aspect or the semantic aspect by having them answer the different types of questions under both types of tasks. Valence of one modality was successfully decoded from the classifiers trained from the other modality, which is a modality-general processing of affect. However, please note that even under the semantic task, the participants paid attention to the stimuli. Thus at least modality-specific valence information even under semantic task could be automatically encoded. However, modality-general may require attention. The semantic task in this study required processing of the video or music to determine its content, which could then have engaged modality-general affective processing. A future study should directly compare the effects of a task that requires a response and no task on the nature of affective processing.

6.2. Merit and contribution

The current work utilized multiple multivariate techniques to analyze the fMRI data to find the modality-general processing of affect. One of the differences between this work and the previous literature was that the current study included neutral conditions. Because the nature of the distribution of valence and arousal in many of the emotion stimuli sets is U-shaped (high arousal for positive and negative and low arousal for neutral), it was hard to disambiguate the effect of valence and arousal when the neutral condition was included in the study. For the three valence categories in the current stimuli set, the arousal levels were equated so that any difference between the three categories can be attributed to the valence, not confounded with arousal. Another contribution is that this work supported the previous findings with dynamic naturalistic stimuli. For example, most of the visual stimuli were static picture stimuli (Baucom et al., 2012; Chikazoe et al., 2014; Shinkareva et al., 2014) or controlled video stimuli (Kim et al., 2015; Peelen et al., 2012). This study validated previous findings that modality-general representation of valence from static controlled stimuli can be applied to the more naturalistic dynamic stimuli, which is more similar to those that people encounter in their everyday lives. Methodologically, this dissertation illustrates the applications of multiple multivariate techniques including MVPA, MDS, searchlight analysis, and RSA to investigate representation of affective states and those that allowed us to examine similarities and differences in the representation of affect across individuals, stimuli, and tasks based on fMRI data. The outcome of this work serves to further our understanding of how the brain represents valence from different modalities.

6.3. Future directions

One of the limitations of this research was control of the low-level features. Though cross-modal classification completely rules out the effect of the low-level features, within-modal classification results may be confounded with the low level features of the stimuli. In the current study there was no difference of low-level features between valence categories, but still there is a possibility of confounding with the features. A more rigorous way to statistically remove the effect of low-level features is regressing the features out when MPSC is computed. For example, Chikazoe et al. (2014) extracted one component from 5 visual features including local contrast, luminance, hue, number of edges, and visual salience using principal component analysis. Thus MPSC data with the low-level feature regressed out will show the result without the effect of the features.

Another limiting aspect of the study is that the task questions need to be revised. Unlike the result from a separate group of participants, the fMRI participants' behavioral performance was low in terms of the overall accuracy. The accuracies between modality by task should be matched (semantic-music: .46, semantic: video: .55, affective music: .53, affective video: .50) because for the semantic task, music trials were harder than video trials. For future studies, the difficulties across experimental conditions should be balanced.

This work used two modalities, visual and auditory, and demonstrated modality-general processing of affect from multiple regions of interests. Searchlight analyses were able to pinpoint the informative voxels to modality-general representation of affect and modality. An fMRI design is a useful tool to answer *where*, whereas it is less helpful to answer *when* questions. An EEG study may be more suitable to address a when question. Specifically, if modality-general processing occurs, then this study does not answer when it happens. For example, Aftanas et al. (2002) found that all affective vs. low arousal IAPS pictures induce a greater amount of the theta synchronization over posterior regions in the early post-stimulus period of 1 s. They also found the larger synchronization between high and low arousal pictures in the 200-700 ms time window, suggesting that picture valence can be identified in a very short time. Another advantage of an EEG study is to disambiguate the timings of modality-specific and modality-general processing. One possibility is that if early parts of the signal can code valence, then that would likely be modality-specific automatic encoding. One could then look for a later wave form that might reflect modality-general processing. Similar methods of cross-modal prediction could be used on EEG data to determine if the processing being examined is modality-specific or modality general. These types of multivariate applications would open up new ways to understand EEG data.

References

Abelson, R. P., & Sermat, V. (1962). Multidimensional scaling of facial expressions. *Journal of Experimental Psychology, 63*, 546-554.

Aftanas, L. I., Varlamov, A. A., Pavlov, S. V., Makhnev, V. P., & Reva, N. V. (2002). Time-dependent cortical asymmetries induced by emotional arousal: EEG analysis of event-related synchronization and desynchronization in individually defined frequency bands. *International Journal of Psychophysiology*, *44*(1), 67-82.

Anders, S., Lotze, M., Erb, M., Grodd, W., & Birbaumer, N. (2004). Brain activity underlying emotional valence and arousal: A response-related fMRI study. *Human Brain Mapping*, *23*(4), 200-209.

Barrett, L.F., & Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Advances in Experimental Social Psychology, 41*, 167-218.

Barrett, L.F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2007). The experience of emotion. *Annual Review of Psychology,* 373–403.

Baucom, L. B., Wedell, D. H., Wang, J., Blitzer, D. N., & Shinkareva, S. V. (2012). Decoding the neural representation of affective states. *Neuroimage*, *59*(1), 718-727.

Bigand, E., Vieillard, S., Madurell, F., Marozeau, J., & Dacquet, A. (2005). Multidimensional scaling of emotional responses to music: The effect of musical expertise and of the duration of the excerpts. *Cognition & Emotion*, *19(8)*, 1113-1139.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Bradley, M. M., & Lang, P. J. (2000). Emotion and motivation. *Handbook of psychophysiology*, *2*, 602-642.

Bradley, M., & Lang, P. J. (1999). *The International affective digitized sounds (IADS)[: stimuli, instruction manual and affective ratings*. NIMH Center for the Study of Emotion and Attention.

Brandmeyer, A., Sadakata, M., Spyrou, L., McQueen, J. M., & Desain, P. (2013). Decoding of single-trial auditory mismatch responses for online perceptual monitoring and neurofeedback. *Frontiers in Neuroscience*, *7*.

Cacioppo, J.T., Tassinary, L.G., & Berntson, G.G. (2000). Psychophysiological science: Interdisciplinary approaches to classic questions about the mind. In J.T. Cacioppo, L.G. Tassinary, & G.G. Berntson, (Eds.), *Handbook of Psychophysiology* (pp. 1-16). Cambridge, England: Cambridge University Press.

Calvo, M. G., Nummenmaa, L., & Avero, P. (2008). Visual search of emotional faces: Eye-movement assessment of component processes. *Experimental Psychology*, *55*(6), 359-370.

Cauchoix, M., Barragan-Jason, G., Serre, T., & Barbeau, E. J. (2014). The neural dynamics of face detection in the wild revealed by MVPA. *The Journal of Neuroscience*, *34*(3), 846-854.

Chavez, R. S., & Heatherton, T. F. (2014). Representational similarity of social and valence information in the medial pFC. *Journal of Cognitive Neuroscience, 27,* 73–82.

Chikazoe, J., Lee, D. H., Kriegeskorte, N., & Anderson, A. K. (2014). Population coding of affect across stimuli, modalities and individuals. *Nature Neuroscience, 17(8),* 1114–1122.

Cohen, M. A., Horowitz, T. S., & Wolfe, J. M. (2009). Auditory recognition memory is inferior to visual recognition memory. *Proceedings of the National Academy of Sciences*, *106*(14), 6008-6010.

Connolly, A. C., Guntupalli, J. S., Gors, J., Hanke, M., Halchenko, Y. O., Wu, Y. C., & Haxby, J. V. (2012). The representation of biological classes in the human brain. *The Journal of Neuroscience, 32(8),* 2608-2618.

Coutanche, M. N., Thompson-Schill, S. L., & Schultz, R. T. (2011). Multi-voxel pattern analysis of fMRI data predicts clinical symptom severity. *Neuroimage, 57(1),* 113-123.

Coutinho, E., & Cangelosi, A. (2011). Musical Emotions: Predicting second-by-second subjective feelings of emotion from low-level psychoacoustic features and physiological measurements. *Emotion, 11(4),* 921–937.

Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) “brain reading”: detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage, 19(2),* 261-270.

Cunningham, W., Raye, C. L., & Johnson, M. (2004). Implicit and explicit evaluation: fMRI correlates of valence, emotional intensity, and control in the processing of attitudes. *Journal of Cognitive Neuroscience*, *16*(10), 1717-1729.

Eerola, T., & Vuoskoski, J. K. (2010). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, 18-49.

Eisenbarth, H., & Alpers, G. W. (2011). Happy mouth and sad eyes: scanning emotional facial expressions. *Emotion*, *11*(4), 860-865.

Ethofer, T., Van De Ville, D., Scherer, K., & Vuilleumier, P. (2009). Decoding of emotional information in voice-sensitive cortices. *Current Biology*, *19*(12), 1028-1033.

Etkin, A., Egner, T., & Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends in Cognitive Sciences*, *15*(2), 85-93.

Etzel, J. A., Zacks, J. M., & Braver, T. S. (2013). Searchlight analysis: promise, pitfalls, and potential. *Neuroimage*, *78*, 261-269.

Fairchild, M.D. (2005). *Color Appearance Models*. John Wiley & Sons Ltd.

Frijda, N. H. (1986). *The emotions*. London, England: Cambridge University Press.

Gable, P. A., & Harmon-Jones, E. (2008a). Approach-motivated positive affect reduces breadth of attention. *Psychological Science, 19*, 476-482.

Gable, P. A., & Harmon-Jones, E. (2008b). Relative left frontal activation to appetitive stimuli: Considering the role of individual differences. *Psychophysiology, 45*, 275-278.

Gable, P., & Harmon-Jones, E. (2010a). The blues broaden, but the nasty narrows: Attentional consequences of negative affects low and high in motivational intensity. *Psychological Science, 21(2),* 211-215.

Gable, P., & Harmon-Jones, E. (2010b). The motivational dimensional model of affect: Implications for breadth of attention, memory, and cognitive categorisation. *Cognition & Emotion, 24(2),* 322-337.

Gabrielsson, A., & Lindström, E. (2010). The role of structure in the musical expression of emotions. In P.N. Juslin & J.A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 367-400).

Habel, U., Klein, M., Kellermann, T., Shah, N. J., & Schneider, F. (2005). Same or different? Neural correlates of happy and sad mood in healthy males. *Neuroimage*, *26*(1), 206-214.

Habes, I., Krall, S. C., Johnston, S. J., Yuen, K. S. L., Healy, D., Goebel, R., & Linden, D. E. J. (2013). Pattern classification of valence in depression. *NeuroImage: clinical*, *2*, 675-683.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*(5539), 2425-2430.

Haynes, J. D., & Rees, G. (2005a). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience*, *8*(5), 686-691.

Haynes, J. D., & Rees, G. (2005b). Predicting the stream of consciousness from activity in human visual cortex. *Current Biology*, *15*(14), 1301-1307.

Haynes, J. D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, *7*(7), 523-534.

Henderson J. Shinkareva S. Wang J. Luke S. Olejarczyk J. (2013). Predicting cognitive state from eye movements. *Plos One, 8* (5), e64937.

Holland, P. C., & Gallagher, M. (1999). Amygdala circuitry in attentional and representational processes. *Trends in Cognitive Sciences*, *3*(2), 65-73.

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science*, *310*(5754), 1680-1683.

Hutcherson, C. A., Goldin, P. R., Ochsner, K. N., Gabrieli, J. D., Barrett, L. F., & Gross, J. J. (2005). Attention and emotion: does rating emotion alter neural responses to amusing and sad films? *Neuroimage*, *27*(3), 656-668.

Joormann, J., & Gotlib, I.H. (2007). Selective attention to emotional faces following recovery from depression. *Journal of Abnormal Psychology, 116,* 80-85.

Juslin, P. N., & Laukka, P. (2004). Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *Journal of New Music Research*, *33*(3), 217-238.

Kassam, K. S., Markey, A. R., Cherkassky, V. L., Loewenstein, G., & Just, M. A. (2013). Identifying Emotions on the Basis of Neural Activation. *PLoS ONE,8*(6), e66032.

Kay, K. N., Rokem, A., Winawer, J., Dougherty, R. F., & Wandell, B. A. (2013). GLMdenoise: a fast, automated technique for denoising task-based fMRI data. *Frontiers in Neuroscience, 7*.

Kennedy, D. P., & Adolphs, R. (2012). Perception of emotions from facial expressions in high-functioning adults with autism. *Neuropsychologia, 50(14)*, 3313-3319.

Kim, J., Schultz, J., Rohe, T., Wallraven, C., Lee, S. W., & Bülthoff, H. H. (2015). Abstract Representations of Associated Emotions in the Human Brain. *The Journal of Neuroscience*, *35*(14), 5655-5663.

Klasen, M., Kenworthy, C. A., Mathiak, K. A., Kircher, T. T. J., & Mathiak, K. (2011). Supramodal representation of emotions. *Journal of Neuroscience, 31*, 13635–13643.

Klasen, M., Kreifelts, B., Chen, Y. H., Seubert, J., & Mathiak, K. (2014). Neural processing of emotion in multimodal settings. *Frontiers in Human Neuroscience*, *8*.

Koelsch, S., Fritz, T., Müller, K., & Friederici, A. D. (2006). Investigating emotion with music: an fMRI study. *Human Brain Mapping*, *27*(3), 239-250.

Kolodyazhniy, V., Kreibig, S.D., Gross, J.J., Roth, W.T., & Wilhelm, F.H. (2011). An affective computing approach to physiological emotion specificity: toward subject-independent and stimulus-independent classification of film-induced emotions. *Psychophysiology, 48(7),* 908-922.

Kosslyn, S. M., Thompson, W. L., Kim, I. J., & Alpert, N. M. (1995). Topographical representations of mental images in primary visual cortex. *Nature*, *378*(6556), 496-498.

Kreibig, S.D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology, 84,* 394–421.

Kreibig, S.D., Samson, A.C., & Gross, J.J. (2010). The psychophysiology of mixed emotional states. *Psychophysiology, 44(5),* 787-806.

Kreibig, S.D., Wilhelm, F.H., Roth, W.T., Gross, J.J. (2007). Cardiovascular, electrodermal, and respiratory response patterns to fear- and sadness-inducing films. *Psychophysiology, 44(5),* 787-806.

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(10), 3863-3868.

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis–connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience, 2*.

Kumpik, D. P., Roberts, H. E., King, A. J., & Bizley, J. K. (2014). Visual sensitivity is a stronger determinant of illusory processes than auditory cue parameters in the sound-induced flash illusion. *Journal of Vision*, *14*(7), 12.

Lakens, D., Fockenberg, D. A., Lemmens, K. P., Ham, J., & Midden, C. J. (2013). Brightness differences influence the evaluation of affective pictures. *Cognition & Emotion*, *27*(7), 1225-1246.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: brain mechanisms and psychophysiology. *Biological Psychiatry*, *44*(12), 1248-1263.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1999). International affective picture system (IAPS): Instruction manual and affective ratings. *The center for research in psychophysiology, University of Florida*.

Lang, P. J., Bradley, M. M., Fitzsimmons, J. R., Cuthbert, B. N., Scott, J. D., Moulder, B., & Nangia, V. (1998). Emotional arousal and activation of the visual cortex: an fMRI analysis. *Psychophysiology*, *35*(2), 199-210.

Lange, K., Williams, L. M., Young, A. W., Bullmore, E. T., Brammer, M. J., Williams, S. C., ... & Phillips, M. L. (2003). Task instructions modulate neural responses to fearful facial expressions. *Biological Psychiatry*, *53*(3), 226-232.

Lartillot, O., Toiviainen, P., & Eerola, T. (2008). A Matlab Toolbox for Music Information Retrieval, in C. Preisach, H. Burkhardt, L. Schmidt-Thieme, R. Decker (Eds.), *Data Analysis, Machine Learning and Applications, Studies in Classification, Data Analysis, and Knowledge Organization*, Springer, Verlag.

Laurier, C., Grivolla, J., & Herrera, P. (2008, December). Multimodal music mood classification using audio and lyrics. In *Machine Learning and Applications, 2008. ICMLA'08. Seventh International Conference on* (pp. 688-693). IEEE.

LeDoux, J. (1998). *The emotional brain: The mysterious underpinnings of emotional life*. Simon and Schuster.

Lindquist, K. A., & Barrett, L. F. (2012). A functional architecture of the human brain: emerging insights from the science of emotion. *Trends in Cognitive Sciences*, *16*(11), 533-540.

Lindquist, K. A., Satpute, A. B., Wager, T. D., Weber, J., & Barrett, L. F. (2015). The brain basis of positive and negative affect: evidence from a meta-analysis of the human neuroimaging literature. *Cerebral Cortex*, 1-13.

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: a meta-analytic review. *Behavioral and Brain Sciences*, *35*(03), 121-143.

Looser, C. E., Guntupalli, J. S., & Wheatley, T. (2013). Multivoxel patterns in face-sensitive temporal regions reveal an encoding schema based on detecting life in a face. *Social Cognitive and Affective Neuroscience*, *8*(7), 799-805.

Matsuda, S., & Yamamoto, J. (2015). Intramodal and cross-modal matching of emotional expression in young children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 10*, 109-115.

Mesulam, M. M. (2000). Behavioral neuroanatomy: large-scale networks, association cortex, frontal syndromes, the limbic system, and hemispheric specializations. *Principles of Behavioral and Cognitive Neurology*, *2*, 1-120.

Meyer, K., Kaplan, J. T., Essex, R., Webber, C., Damasio, H., & Damasio, A. (2010). Predicting visual stimuli on the basis of activity in auditory cortices. *Nature Neuroscience*, *13*(6), 667-668.

Mothes-Lasch, M., Mentzel, H. J., Miltner, W. H., & Straube, T. (2011). Visual attention modulates brain activation to angry voices. *The Journal of Neuroscience*, *31*(26), 9594-9598.

Omar, R., Henley, S. M., Bartlett, J. W., Hailstone, J. C., Gordon, E., Sauter, D. A., ... & Warren, J. D. (2011). The structural neuroanatomy of music emotion recognition: evidence from frontotemporal lobar degeneration. *Neuroimage, 56*(3), 1814-1821.

Onton, J., & Makeig, S. (2009). High-frequency broadband modulations of electroencephalographic spectra. *Frontiers in Human Neuroscience*, *3*.

Oosterhof, N. N., Wiggett, A. J., Diedrichsen, J., Tipper, S. P., & Downing, P. E. (2010). Surface-based information mapping reveals crossmodal vision–action representations in human parietal and occipitotemporal cortex. *Journal of Neurophysiology*, *104*(2), 1077-1089.

Peelen, M. V., Atkinson, A. P., & Vuilleumier, P. (2010). Supramodal representations of perceived emotions in the human brain. *The Journal of Neuroscience, 30*(30), 10127-10134.

Pereira, F., & Botvinick, M. (2011). Information mapping with pattern classifiers: a comparative study. *Neuroimage*, *56*(2), 476-496.

Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage, 45*(1), S199-S209.

Pessoa, L., McKenna, M., Gutierrez, E., & Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences, 99*(17), 11458-11463.

Philip, R. C. M., Whalley, H. C., Stanfield, A. C., Sprengelmeyer, R., Santos, I. M., Young, A. W., & Hall, J. (2010). Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychological medicine, 40(11),* 1919-1929.

Posner, J., Russell, J.A., & Peterson, B.S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Developmental Psychopathology, 17(3),* 715–734.

Pye, A., & Bestelmeyer, P. E. (2015). Evidence for a supra-modal representation of emotion from cross-modal adaptation. *Cognition*, *134,* 245-251.

Roberts, J. S., & Wedell, D. H. (1994). Context effects on similarity judgments of multidimensional stimuli: Inferring the structure of the emotion space. *Journal of Experimental Social Psychology*, *30*(1), 1-38.

Roseman, I.J., & Smith, C.A. (2001). Appraisal theory: overview, assumptions, varieties, controversies. In Roseman, Ira J.; Smith, Craig A. Scherer, Klaus R. (Ed); Schorr, Angela (Ed); Johnstone, Tom (Ed), *Appraisal processes in emotion: Theory, methods, research. Series in affective science* (pp. 3-19). New York, NY: Oxford University Press.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145.

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of Personality and Social Psychology*, *76*(5), 805.

Russell, J.A. & Barrett, L.F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology, 76(5),* 805-819.

Russell, J. A., & Bullock, M. (1985). Multidimensional scaling of emotional facial expressions: similarity from preschoolers to adults. *Journal of Personality and*

Social Psychology, 48(5), 1290.

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect grid: a single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology, 57(3)*, 493-502.

Russell, J.A. (1980a). A circumplex model of affect. *Journal of Personality and Social Psychology, 39(6),* 1161-1178.

Russell, J.A. (1980b). Emotion, core affect, and psychological construction. *Cognition & Emotion*, *23(7),* 1259-1283.

Russell, J.A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110(1),* 145-172.

Russell, J.A., & Carroll, J.M. (1999). On the bipolarity of positive and negative affect. *Psychology Bulletin, 125*, 3–30.

Russo, F. A., Vempala, N. N., & Sandstrom, G. M. (2013). Predicting musically induced emotions from physiological inputs: linear and neural network models. *Frontiers in Psychology*, *4*.

Said, C. P., Moore, C. D., Engell, A. D., Todorov, A., & Haxby, J. V. (2010). Distributed representations of dynamic facial expressions in the superior temporal sulcus. *Journal of Vision*, *10*(5), 11.

Schaefer, R. S., Farquhar, J., Blokland, Y., Sadakata, M., & Desain, P. (2011). Name that tune: decoding music from the listening brain. *Neuroimage*, *56*(2), 843-849.

Scherer, K. R. (1999). Appraisal theories. In T. Dalgleish, & M. Power (Eds.). *Handbook of Cognition and Emotion (pp. 637–663)*. Chichester: Wiley.

Sestito, M., Umiltà, M. A., De Paola, G., Fortunati, R., Raballo, A., Leuci, E., & Gallese, V. (2013). Facial reactions in response to dynamic emotional stimuli in different modalities in patients suffering from schizophrenia: a behavioral and EMG study. *Frontiers in Human Neuroscience*, *7*.

Shinkareva, S. V., Kim, J., Wang, J., & Wedell, D. H. (in preparation). Decoding affective states from audiovisual stimuli.

Shinkareva, S. V., Wang, J., & Wedell, D. H. (2013). Examining similarity structure: multidimensional scaling and related approaches in neuroimaging. *Computational and Mathematical Methods in Medicine*, *2013*.

Shinkareva, S. V., Wang, J., Kim, J., Facciani, M. J., Baucom, L. B., & Wedell, D. H. (2014). Representations of modality-specific affective processing for visual and auditory stimuli derived from functional magnetic resonance imaging data. *Human Brain Mapping*, *35*(7), 3558-3568.

Smith, C. A., & Ellsworth P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology, 48(4),* 813-838.

Straube, T., Kolassa, I. T., Glauer, M., Mentzel, H. J., & Miltner, W. H. (2004). Effect of task conditions on brain responses to threatening faces in social phobics: an event-related functional magnetic resonance imaging study. *Biological Psychiatry*, *56*(12), 921-930.

Surguladze, S., Brammer, M. J., Keedwell, P., Giampietro, V., Young, A. W., Travis, M. J., & Phillips, M. L. (2005). A differential pattern of neural response toward sad versus happy facial expressions in major depressive disorder. *Biological psychiatry, 57(3)*, 201-209.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, *315*(5811), 515-518.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., ... & Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, *15*(1), 273-289.

Vuilleumier, P., & Schwartz, S. (2001). Emotional facial expressions capture attention. *Neurology*, *56*(2), 153-158.

Wager, T. D., Davidson, M. L., Hughes, B. L., Lindquist, M. A., & Ochsner, K. N. (2008). Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron*, *59*(6), 1037-1050.

Wager, T. D., Lindquist, M., & Kaplan, L. (2007). Meta-analysis of functional neuroimaging data: current and future directions. *Social Cognitive and Affective Neuroscience*, *2*(2), 150-158.

Watson, R., Latinus, M., Noguchi, T., Garrod, O., Crabbe, F., & Belin, P. (2014). Crossmodal adaptation in right posterior superior temporal sulcus during face–voice emotional integration. *The Journal of Neuroscience*, *34*(20), 6813-6821.

Webster, G. D., & Weir, C. G. (2005). Emotional responses to music: Interactive effects of mode, texture, and tempo. *Motivation and Emotion*, *29*(1), 19-39.

Whalen, P. J., Kagan, J., Cook, R. G., Davis, F. C., Kim, H., Polis, S., & Johnstone, T. (2004). Human amygdala responsivity to masked fearful eye whites. *Science*, *306*(5704), 2061-2061.

Wilson-Mendenhall, C., Barrett, L. F., & Barsalou, L. W. (2013). Neural evidence that human emotions share core affective properties. *Psychological Science, 24*, 947-956.

APPENDIX A – INSTRUCTIONS FOR FMRI SESSIONS

1) Video affective session:

This session is Video Judgment.

All of the trials are video and all of the questions are about emotion.

The questions are "Positive?", "Negative?", “Neutral?"

You will always respond with

INDEX finger for YES

MIDDLE finger for NO

The trial will begin soon.

2) Video semantic session:

This session is Video Object.

All of the trials are video clips and all of the questions are about the object in the video.

The questions are "Human?", "Animal?", “Scene?"

You will always respond with

INDEX finger for YES

MIDDLE finger for NO

The trial will begin soon.

3) Music affective session:

This session is Music Judgment.

All of the trials are musical clips and all of the questions are about emotion.

The questions are "Positive?", "Negative?", “Neutral?"

You will always respond with

INDEX finger for YES

MIDDLE finger for NO

The trial will begin soon.

4) Music semantic session:

This session is Music Instrument.

All of the trials are musical clips and all of the questions are about the instruments.

The questions are "String?", "Wind?", “Percussion?"

You will always respond with

INDEX finger for YES

MIDDLE finger for NO

The trial will begin soon.