***If You See What I See, Then I know How You Feel*: HOW SCENE CONTEXT AND FACIAL MIMICRY DIFFERENTIALLY AFFECT SITUATED RECOGNITION OF DYNAMIC AND STATIC EXPRESSIONS**

by

Cindy Hamon-Hill

Submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
August 2016

TABLE OF CONTENTS

LIST OF TABLES

List of Figures

Abstract

Testing facial expression recognition has primarily relied on facial stimuli consisting of static images of posed expressions or computer generated animations. In recent years, the ecological validity of results based on such stimuli has come under scrutiny. It was important for the research in this thesis to increase ecological validity by testing expression recognition to natural dynamic facial expressions presented with a context scene. Accounting for an effect of context on expression recognition was one objective in the current work. Another objective was to test the explanation for expression recognition offered by simulation theory involving facial mimicry. Past studies demonstrated that a disruption to mimicry impaired expression recognition in static and computer animated faces. Disrupted mimicry had not been tested on natural, dynamic expressions presented with scene context. In the task for this research participants were presented with multiple pairings of a natural dynamic facial expression and a context scene, and asked to judge whether the model's expression was in response to the given context scene. It was hypothesized that a disruption to facial mimicry that was intended to interfere with motor activity in the simulation process would impair recognition to static more than dynamic expressions. It was also hypothesized that when expressions were difficult to discriminate as the true or false response to a context scene, the observer's embodied response to the context scene would bias judgments about the expression as a match to their own.

Over three experiments, the effects of context were always greater on judgments to dynamic than static expressions, and greater on long than brief expressions. Accuracy was better to dynamic long than brief and static expressions. Effects of disrupted mimicry interacted with duration rather than static and dynamic quality of expressions, resulting in a weak facilitation in recognition of long expressions. Overall, the results indicate that static and dynamic expressions are processed differently, that mimicry may not necessarily underlie expression recognition when expressions are presented with context, and that the use of natural dynamic expressions in context may provide a more appropriate paradigm for research into ecologically relevant emotion recognition processes.

Acknowledgments

There are thanks to be expressed to many people for their ongoing support of my work on this thesis. I thank Dr. Patricia McMullen for stepping in, encouraging me, and seeing me through to completion. I also thank Dr. Gail Eskes for joining my committee at the 11th hour and particularly for her kind and patient words of encouragement. I thank Dr. John Barresi for his constant support over many years, including times when he should have been enjoying his retirement, or at least picking mushrooms. So much of what we shared and explored in our research doesn't appear on the pages within, but endures and will no doubt appear in between the lines in work to follow. I am so very fortunate to have an abundance of good friends who have always provided understanding and support. Friendships with Geraldine, Sandie, Yoko and Val make a huge difference in my life. Most of all, the completion of this thesis and my graduate work in it's entirety was not possible without Erick's unwavering love and support. Where would I be without him? Definitely not here. I am forever grateful. I think it's quite fitting to say that words cannot adequately express my thanks, but perhaps you can tell by the smile on my face and the sparkle in my eye how I truly feel.

CHAPTER 1: INTRODUCTION

1.1 Situated expression recognition

Successful social interaction involves the ability to assess, understand and respond appropriately to others' beliefs, desires, intentions, and feelings, all of which are examples of mental states that underlie and motivate behavior. Correct assessment of mental states can be challenging because they are not directly observable. To aid in assessment, one looks for non-verbal cues revealed in observable behavior. Importantly, the overt behavior is attributed to that person's inferred mental state. The ability to rapidly process non-verbal cues is especially important for smooth social interaction as it allows the observer to prepare his or her own behavioral response and form predictions regarding subsequent behavior (Goldman, 2009; Malle, 2005).

One source of non-verbal information guiding our judgments about what another person is feeling is facial expressions. Recognition of a facial expression typically involves multiple stages of processing, beginning with early, automatic detection of visual information that supports recognition of the stimulus as a face, followed by recognition of a particular configuration of facial features[1] as an expression signaling a particular emotion (Adolphs, 2002; Niedenthal et al., 2010). According to de Gelder (2008), detection and encoding of a stimulus as a face occurs at approximately 70 ms after stimulus onset, and detection of the stimulus as an expression occurs at approximately 140 ms. This early detection is rapidly followed by semantic processing of the facial expression. To simplify the recognition process in this linear way is misleading. Research

---

[1] In this work, "features" refers to the individual features of a face such as the eyes, brows, nose and mouth, and a "configuration of features" refers to a spatial pattern of the features that varies in different expressions. For example, the spatial patterns of eyebrows vary across expressions of surprise and anger.

on the perception of facial expressions suggests that numerous events, related to processing of factors such as sex, age, identity, and body posture, take place between recognition of the stimulus as a face and semantic processing of that face (Azevier, 2008; Hassein et al., 2013; Righart & de Gelder, 2007). In natural, real-world circumstances, recognition also includes comprehension of the expression as a response to events in the surrounding environment (Carroll & Russell, 1996). In other words, the early perceptual processes that support recognition involve the integration of various forms of sensory information from the facial stimulus itself and contextual information surrounding the face and in that way has been described as a constructive process (Barrett, 2006).

Integration of contextual information need not be modeled as a linear process restricted to semantic processes that are associated with higher order cognitive processing and occurring separately from and substantially later than the visual processing of facial features. Indeed, as will be reviewed, there is ample evidence for an effect of context on early visual processing of facial features. This research began with the realization that conceptual models for recognition of facial expression that do not account for contextual factors in expression processing are limited in their generalizability. Expressions presented in isolation are not representative of daily social experience. Such models provide valuable explanation for very specific aspects of visual processing, such as detection the configural organization of features, and whether an expression is a smile or frown. However, that the same mechanisms supporting feature detection would also support judgments about an expression as appropriate to a situation was questionable. Models of the recognition of facial expressions that integrate contextual information into the processing of facial features may offer greater generalizability and enable a richer understanding of the mechanisms that guide social behavior.

This thesis investigates the role of situational context in expression recognition. Situational context refers to a given event or set of circumstances that imply a purpose or expectations about behavior (such as a wedding, funeral or romantic date). Following from this, the recognition process that integrates details of feature configuration with situational information to support inferences about the expression is referred in this thesis as situated facial expression recognition. Situated facial expression recognition best represents how one perceives facial expressions in the natural world – not in isolation but rather within a social context that conveys particular details about a given situation that may or may not include other people. Therefore, situated expression recognition is argued to be ecologically relevant for understanding processes that support recognition of the emotional state of another individual and social interaction.

1.2 Theoretical perspectives

There are two major theoretical perspectives explaining how individuals understand the mental state of others based on facial expression recognition and other explicit cues. One is the ‘theory-theory’ perspective of theory of mind and the other is simulation theory. Theory-theory suggests that individuals can infer what another individual is feeling by attributing that individual’s observable behavior (e.g., facial expressions) to his or her unobservable mental states. According to this perspective, underlying processes that support recognition involve the formation of naïve theories (Gopnik & Wellman, 1992; Saxe, 2005). This process entails hypothetical reasoning about the likelihood of another individual’s response to a relevant situation based on conceptual knowledge, beliefs, and expectations regarding both the individual and the situation. In other words, based on their beliefs and knowledge, an observer develops theories about causal relationships between social situations, motivations, and overt behaviour (emotional expressions and actions). The observer uses these theories to infer the meaning and attribute underlying motivations and emotions to the behavior (Adolphs, 2002; Gopnik & Wellman, 1994; Malle, 2005; Saxe, 2005). In doing so, individuals look for evidence to support existing theories and, when external evidence is inconsistent with the existing theory, that theory is modified. This process of refinement continues over an individual's lifetime, as they accumulate conceptual knowledge about the behaviour of others.

Critics of the theory-theory perspective suggest the inferential process associated with interpreting intention, motivation or feeling from an expression is not always necessary. Although such elaborate hypothetical processing may be necessary at given times (e.g., when the person or situation is unfamiliar), the time and demand on explicit cognitive processing would be otherwise costly and likely impair smooth social interaction (Goldman, 2005; Iacoboni, 2008).

The second major perspective, known as simulation theory, suggests that judgments about the emotional states of another individual are based on the observer's reenacted affective states of the other individual. In order to understand what another person is thinking or feeling, the observer reenacts or generates by pretense the same emotional event as the person under observation or consideration. The simulated experience may be deliberate and supported by explicit cognitive process. For example, consciously adopting the perspective of another individual and imagining their experience in a given situation as one's own, leads to an understanding of the emotional experience. However, another version of simulation theory emphasizes automaticity, with the simulated experience occurring without conscious intention (Gallese & Goldman, 1998). Most importantly, in contrast to the linear, compartmentalized approach of theory-theory, simulation theory proposes that perception and simulation of a facial expression is a constructive process that functions as a dynamic, multi-system response to perceived events, with perception, action, and activation of mental structures comprising a unified process in the brain (Barrett, 2014; Iacoboni, 2008; 2009). Compared to theory-theory, which relies on concepts and logical, hypothetical reasoning, simulation theory suggests a process that is rapid, experientially based and less demanding in terms of cognitive resources. Recognition of another person's emotional state by means of simulating the perceived state has provided a compelling framework for investigating the processes underlying facial expression recognition.

A key point about simulation theory as it relates to expression recognition is correspondence between the perception and production of a facial expression for a particular emotion. Specifically, models of face-based emotion recognition have proposed that perception of the emotional state of another individual automatically activates a simulation of the expression and corresponding internal state in the self. The simulated state represents the meaning of the expression that is automatically attributed to that other person (Goldman & Sripada, 2005; Goldman, 2009). In a "Simulation of Smiles" (SIMS) model of expression recognition, perception of a smile is followed by a simulation of the affective state that corresponds with the smile, described as activation of the brain's motor, somatosensory, affective, and reward systems (Niedenthal et al., 2010). Correspondence has been supported by evidence of common neural activation associated with perceiving and experiencing a given emotion. For example, perceiving an expression of disgust and smelling noxious stimulants activated common nuclei in the insula, and lesions to the insula were associated with impaired recognition of another person expressing disgust as well as experiencing disgust to a noxious stimulant (Jabbi & Keysers, 2008: Wicker et al, 2003). Similarly, activation of the amygdala has been commonly associated with perception and expression of high arousal emotions such as fear, or perceived threat and individuals with lesioned amydalae have demonstrated lower levels of experienced fear or arousal at a threat and lower ratings of stimuli as threatening compared to controls (Adophs, 2002).

A central role for simulation theory in facial expression recognition is supported by empirical research on facial mimicry, wherein an observer spontaneously, involuntarily mimics the perceived expression of another person. For instance, perceived fear in an expression triggers automatic mimicry of the expression and a simulated state of fear. The emotion experienced associated with the expression (e.g., fear) is not confused as being one's own; rather, it is automatically recognized as the state of the other individual (Goldman 2005, 2009, 2011; Niedenthal, 2007). Mimicry is considered a key component in the simulation process underlying expression recognition (Iacoboni, 2008), and has been proposed as a mechanism to facilitate an observer's recognition of the emotional state of another person (Atkinson & Adolphs, 2005; Iacoboni, 2008; Gallese, 2003; Goldman & Sripada, 2005; Maringer, Fischer, Krumhuber, & Niedenthal, 2011; Niedenthal et al., 2001,2010; Oberman et al, 2007; Ponari et al, 2012; Rychlowska et al, 2014).

Critics of simulation theory suggests there is no accountability for error – that simulations based on the other person's experience should always lead to a correct interpretation, yet errors in judgment about the intentions of others are common (Saxe, 2005). An important component of the simulation process that is not accounted for in this particular criticism, but is of great interest to this thesis, is the subjectivity of the simulation process. Rarely has there been accountability for the role of the observer in expression recognition. Recognition and judgments about the meaning of another person's expression involve more than processing the signal in the content of perceived facial expression. Processes that directly involve the observer have seldom been accounted for. One person's simulation of another person's observable behavior is a dynamic response, one that is vulnerable to multiple externalized and internalized factors that shape the simulated experience. An observer's gender, state of physiological arousal, or relation to the other person have been shown to affect perception and judgments (Elfenbein & Ambady, 2002; Hess & Fischer, 2014). An imminent externalized factor is a shared situational context. The simulator's response to a shared situation was expected to affect perception and judgments that would influence expression recognition. This was a focal point of interest for this thesis.

1.3 Objectives

A main objective for this thesis was to test the explanation for expression recognition offered by simulation theory; that an observer simulates the perceived state of another individual and attributes the generated state not to him or her self but to the other. Facial mimicry has been identified as an important mechanism underlying recognition by means of contributing to the simulation process (Gallese, 2003, Goldman & Sripada, 2005; Iacoboni, 2008). In this thesis, it was questioned if mimicry underlies recognition when expressions are situated. It was also questioned if recognition of another person's facial expression is influenced by his/her own response to a shared situational context. This possibility suggests that a response to a given context primes perception of the facial expression, thereby influencing judgment about the underlying emotional state of the other person. Findings about the impact of mimicry and/or the observer's response to situational context would support a greater understanding of the simulation processes for recognition of situated expressions.

1.4 Overview

In this thesis, Chapter 2 contains a review of past research on facial expression recognition and discusses the limitations of those findings. It also explores the inclusion and influence of context within expression recognition research, as well as a need for natural facial stimuli in expression recognition paradigms. Lastly, the chapter describes previous research regarding the role of mimicry in expression recognition studies.

Chapter 3 outlines the methodological paradigm that was developed for the research. Chapters 4 through 6 present the experiments conducted. The first outlines an experiment that addressed the research questions pertaining to mimicry and context. Two manipulations previously applied in past research that were intended to disrupt mimicry were implemented during the situated recognition task. Due to the surprising results in Chapter 4 of a facilitation effect of disrupted mimicry, which was opposite to reports in literature, Chapters 5 and 6 explored the effects of mimicry and context on dynamic and static qualities of facial expressions and the duration of their presentation on situated recognition. The results are discussed in Chapter 7.

Chapter 2: Literature Review

2.1 Early research in expression recognition

Research on emotion recognition has been strongly influenced by the writings of Darwin (1872/1998), who proposed that facial expressions, defined as specific patterns of particular facial muscles, are externalized signals of basic biological emotional states. For instance, sneering with an open mouth, curled lips, and exposed teeth signal anger, whereas downturned corners of the mouth signal suffering and sadness, and upturned corners of the mouth signal pleasure (Darwin, 1872/1965). Darwin's idea of specific facial expressions providing signals that reliably communicate discrete emotional states was subsequently supported by abundant research (Tomkins, 1962; Ekman, 1972; Izard, 1971). Evolutionary theorists propose that expressions are adaptive because they successfully communicate one's intentions to the extent that cohorts attend to expressions and respond accordingly, making expressions an efficient tool for navigating social interactions (Fridlund,1991;1994).

Consistent with the evolutionary perspective, faces were for decades considered the purest expression of emotion because they contained core signals, with few cultural differences found regarding the expression of specific emotions (Ekman, Friesen, & Ellsworth, 1972; Ekman, 1993; Izard, 1971, 1994; Tomkins, 1962).[2] Consequently, these emotions were identified as basic, universal biological states, and specific facial

---

[2] Universality pertains to basic expressions and not to more specific expressions (such as jealousy). However, debate continues about basic emotions as well due to reports of an "in-group advantage". Evidence for such an advantage exists in findings of more accurate recognition of basic emotions expressed by members of one's own cultural in-group (Elfenbein & Ambady, 2002). Marsh and Ambady (2008) suggested that the differences in accuracy could be explained by cultural variants, whereby, facial expressions, like language, contain "non-verbal accents" which would change the fine-grain details of an expression.

expressions were considered unambiguous signals of those states (Ekman & Friesen, 1972). These basic emotions were fear, anger, disgust, sadness, surprise, and, sometimes joy.

Historically, studies in emotion recognition primarily measured the ability to correctly categorize facial expressions as one of the basic emotions. This early emphasis on categorization provided valuable information about the processing of relevant perceptual cues and identification of underlying mechanisms that support rapid recognition of emotion in facial features. A standardized set of stimuli representing basic emotions was generated for testing categorization (Ekman and Friesen, 1976), and consisted of non-natural static images of exaggerated facial expressions posed by actors intentionally generating specific facial actions containing the key features of the basic emotions. Expression prototypes meet criteria for each of the basic emotions defined by a facial action coding system (FACS) whereby activation of specific facial muscles is measured in facial action units (Ekman & Friesen, 1976, See Figure 2.1). Though static, the photographs were intended to provide sufficient information to allow an observer to instantly discriminate one of the basic emotions from another. These stimuli have had an enormous impact on decades of emotion recognition research.

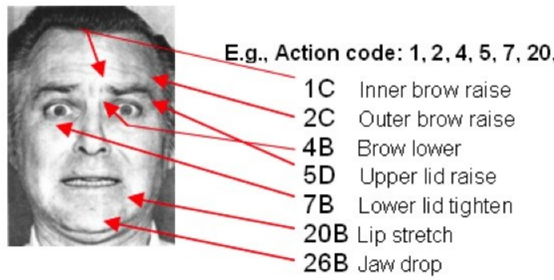

Figure 2.1 Example of Facial Action Coding units (FACS) for the expression fear

Presentation of non-natural facial stimuli became increasingly common as testing stimuli. Static, exaggerated expressions that were, posed, isolated and disembodied became the prototypes for representing a single basic emotion. These non-natural expressions were the dominant stimulus in studies that focused solely on key features of an expression and tested categorization of emotions. Over time, testing materials evolved so that a facial display could be manipulated to display variation in the expression and create the appearance of change in a dynamic expression. For instance, computer generated animations were built from overlaid static images that were “morphed” to generate single images representing different stages in an expression (e.g., from neutral to full expression) and changes in expression (e.g., from positive to neutral then negative). A “bubble” technique was later developed to selectively test recognition or “decoding” of key features and spatial frequencies of an expression that signal each basic emotion (Gosslyn & Schyns, 2001, in Smith et al., 2013). This technique isolates and makes visible the feature(s) of interest and filters out surrounding regions of the face. For instance, an isolated presentation of the eyes has been used to test whether encoding only the eyes is adequate for recognizing fear (Adolphs, 2006; Smith et al., 2013). Similarly, Baron-Cohen's “Reading the Mind in the Eyes” testing stimuli provide black & white images of only eyes intended to represent non-basic emotions (RMET-Revised, Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). Techniques that isolate selective features continue to be based on the notion that certain perceptual signals contain the minimal information necessary to categorize a particular emotion expression (Ekman, Friesen, & Ellsworth, 1972).

As testing stimuli evolved, increasingly non-naturalistic faces were used to represent mostly basic emotions. Working from expression prototypes, computer generated faces were successfully morphed to intentionally represent specific expressions across age, sex, and cultures. Manipulations in the morphing controlled for degrees of emotional intensity and changing emotions presented in facial stimuli (Niedenthal et al., 2001). Although the stimuli offered good control, the parameters of testing expression recognition seldom went beyond simple categorization of a basic emotion to test for comprehension of an expression as a response to a situation. Detection of specific signals in selective features was sufficient to support categorization of expressions photographs and/or artificially generated facial expressions, however it was uncertain if the same detection processes would support accurate recognition of natural facial expressions that were in response to a situational context.

Variability in expression beyond basic prototypes

More recent perspectives on facial expression perception propose that expressions convey not just basic emotions but also complex ones such as compassion, pride, embarrassment, love, envy, jealousy, and shame (Baron-Cohen et al., 2001; Russell, 1994; Russell, Bachorowski, & Fernandez-Dols, 2003). These emotions are considered complex because they are associated with social interaction, with such emotional responses (e.g., jealousy) attributable to particular social situations involving one's self and others. Not surprisingly, complex emotions are often less discrete and therefore less readily recognizable than basic emotions. Consequently, recognition of a complex emotion in a facial expression may require the support of additional clues, such as contextual information (Barrett, Lindquist, & Gendron, 2007; Carroll & Russell, 1996; Russell & Fernandez-Dols, 1997). Underlining the significance of context, emotional responses generated by actors in a movie that did not adhere to prototypical configurations were attributed to the situation in the scene (Carroll & Russell, 1996). Unlike the standard testing procedure of presenting a disembodied, prototypical face, the inclusion of context provides the viewer with information about what caused the expression they are judging.

2.2 The importance of context in facial expression recognition

Based on the premise that facial expression recognition relies predominantly on bottom-up processing of visual signals from the face itself, presenting a face stimulus in isolation was the standard procedure over most of the history of this research area (as reviewed by Fernandez-Dols and Carroll, 1997). Nonetheless, there were some early studies in which facial stimuli were preceded by contextual information presented in pictures of real-life situations (Munn, 1940, in Cowie, 2010), film sequences (Goldberg, 1951, in Cowie, 2010) or verbal narratives (Goodenough & Tinker, 1931).

There are a number of reasons that context is a point of interest in this thesis. In particular, one goal was to increase the ecological validity of empirical investigations into facial expression recognition by incorporating a contextual element, which has been shown by many studies to influence such recognition (Righart & de Gelder, 2006; 2008; see also reviews by Barrett, Mesquita and Gendron, 2011 and Fernandez-Dols & Carroll, 1997). The integration of context expands expression recognition from a rapid, bottom-up processing of visual signals in a face to one that involves additional, more complex and malleable systems that include top-down processing. The suggestion that context may influence perceptual processing of facial expressions has been supported empirically (Aziever et al, 2008, 2011; Halberstadt, Winkielman, Niedenthal, & Dalle, 2009; Hassin et al., 2013; Kim et al., 2004; Mereen, van Heijnsbergen, & de Gelder, 2005; Niedenthal, Brauer, Halberstadt, Innes-Ker, 2001; Righart & de Gelder, 2006; 2008). Since it provides additional information that guides one's comprehension of the full meaning of a facial expression and can influence an observer's expectations and judgments about the emotional state of the other person as a response to a given situation, the inclusion of context into a recognition paradigm should improve the ecological validity of laboratory tasks. The mechanism(s) by which context guides perception of the facial expression and interacts with other processes, such as spontaneous mimicry, represents another point of interest in the current work.

There is a striking array of evidence illustrating the systematic influence of context on recognition of facial expression, much of it based on manipulating contextual information. Such context manipulation has been done in a number of different ways, including embedding the face on whole body postures (Aviezer et al., 2008; 2011; Hassin, Aviezer, & Bentin, 2013; Mereen, van Heijnsbergen, & de Gelder, 2005), embedding a face into social scenes (Righart and de Gelder, 2006; 2008) and pairing faces with elements such as body gestures and voice (see review by de Gelder et al. 2006), surrounding faces (Masuda et al., 2008), emotion words (Barrett, Lindguist & Gendron, 2007; Halberstadt, Winkielman, Niedenthal & Dalle, 2009; Niedenthal, Brauer, Halberstadt, & Innes-Ker, 2001), and verbal narratives (Goodenough & Tinker, 1931). Such context manipulations have repeatedly demonstrated that information from outside the face and not just the facial features is rapidly incorporated into recognition judgments. Some specific examples of these manipulations are detailed in the following paragraphs because of their particular contributions to this thesis's rationale for including context in an ecologically valid expression recognition experiments.

When images of facial expressions were presented with images of an emotionally congruent or incongruent body posture and gesture, a Stroop-like interference effect was observed. Observers in three different studies were faster and more accurate on judgments about facial expressions when the expressions were paired with a body that matched the expression than when they were paired with a body that did not (Aziever et al, 2008; Hassin et al., 2013; Mereen, van Heijnsbergen, & de Gelder, 2005). When photographs of separate faces and bodies expressing fear and anger were used to create incongruent composites (e.g., a face expressing fear pasted onto a body expressing anger), participants' judgments of the facial expression were biased toward the emotion expressed in the body (Mereen, Heijnsbergen & Gelder, 2005). Considering that participants in that study were explicitly instructed to attend only to the face, Mereen and colleagues concluded that the influence of congruent body context on judgment about the facial expression was rapid and automatic. This view was supported by event related potentials (ERPs) indicating more occipital activity to incongruent relative to congruent face and body composites, with the difference in activity appearing as early as 115 ms after stimulus onset. Not only did the results support their hypothesis for rapid neural sensitivity to context information presented with the facial expression, they also suggested slightly prolonged processing due to incongruence (Mereen, van Heijnsbergen, & de Gelder, 2005).

In a series of analogous behavioral experiments, embedding faces onto emotionally congruent images of bodies increased recognition accuracy to facial expressions that are often difficult to distinguish by seemingly reducing ambiguity of facial expressions (Aviezer et al., 2008). For example, disgust and anger are similar emotions and therefore more difficult to discriminate than disgust and fear or sadness (Ekman & Friesen, 1976; Susskind et al., 2007). Cut-out facial expressions of disgust were paired with cut-out bodies intended to represent disgust, anger, sadness and fear. Participants, instructed to attend to the facial expression, were most accurate in categorizing facial expressions as disgust when the expressions were paired with a congruent body expressing disgust. They were least accurate when the expression was paired with an incongruent angry body: in this case, 87% of responses categorized the disgust expression as anger (Aviezer, 2008). The researchers concluded that congruence reduced the ambiguity of facial expressions and, more specifically, that contextual information from the body influenced the perceptual processing of the face and thereby reduced ambiguity of facial expressions (Aziever et al, 2008).

Because the processing of body context altered early perception, it has been suggested processing context is not necessarily part of a later “post-perceptual interpretive” system (Hassin et al, 2013). Eye-tracking patterns by participants viewing the same face-body composites supported this suggestion (Aziever et al, 2008). Typically, individuals attend to the eye region of a face when processing an angry expression and attend to the mouth region when processing a disgust expression (Ekman, 1992). When presented with congruent or incongruent face-body stimuli during perception of an angry facial expression, fixations in the eye region were faster and greater in number when the facial expression was paired with a congruent angry body than with an incongruent disgust body. Moreover, during perception of a disgust expression, fixations on the mouth were faster when paired with a congruent disgust body than with an incongruent angry body. When the disgust expression was paired with an angry body, fixations were faster and greater in number to the eye region. These eye-tracking patterns revealed that initial fixations on the face were influenced by body context, and that a change in scanning pattern of the face occurred with a change in body context (Aziever et al, 2008). The shift in activity with a change of context suggests that reading expressions is not as simple as the basic emotions perspective suggested (that recognition of expressions, of even basic emotions such as anger and disgust, is a straightforward decoding of invariant signals without incorporation of additional information). Instead, recognition of facial expressions was shown to be under the influence of the early, integrative processing of body-context information (Hassin et al., 2013).

Similar results were found when social context was manipulated by surrounding a central cartoon target face with additional cartoon faces displaying expressions that were either congruent or incongruent with the central target's expression (Masuda, et al., 2008). Eye-tracking measures revealed that a shift in participants' attention occurred at 1000 ms, with scanning patterns moving from the initial fixation location at the centre of the face to the surrounding characters. The researchers suggested that the shift in attention indicated surrounding social context information was incorporated early into the perceptual processing of the facial expression, rather than as part of a later, cognitive process influenced by cultural rules. Ratings of the intensity of emotion in the central cartoon character's expression were greater when the surrounding expressions were congruent.

In addition to further supporting the importance of context, the results suggest that the integration of context into judgments about individual expressions has a cultural dimension: judgments were reportedly influenced by the surrounding facial context more so for Japanese participants than American ones. The results suggested that an expression presented in isolation provides incomplete information for inferring an emotion for Japanese but not necessarily Western viewers (Masuda et al., 2008)[3]; the authors even suggested that American participants did not integrate context into their perception of the facial expressions. However, the preceding reports by Hassin et al., (2013) and Azevier et al. (2008), as well as studies that will be discussed below, suggest that, although the western perceptual style is less holistic, social context is integrated into perception of facial expressions to some degree.

Similar to effects of contextual information related to body and social situation, presentations of emotion words (conceptual context) accompanying facial expressions influenced perception, encoding and recall of ambiguous facial expressions. For instance, morphed facial expressions depicting an equal blend of emotions, such as angry and happy, were encoded as angry when presented along with the word “angry” and happy when presented with the word “happy” (Barrett, Lindguist & Gendron, 2007). This effect is relatively long lasting: as such pairing done during encoding influenced recall at a later time point in the same way in other studies (Halberstadt, Winkielman, Niedenthal & Dalle, 2009; Niedenthal, Brauer, Halberstadt, & Innes-Ker, 2001).

Figure 2.2 illustrates the impact of presenting a face in isolation or with situational information. Not only is it challenging to ascribe a single label to non-prototypical and possibly ambiguous expression on an isolated face, it is not always realistic given the propensity of blended emotions in non-prototypical expressions. Once the face is situated, it is less difficult to recognize the possible blend of emotions such as sorrow,

[3] The ratio of attention to the central character relative to the surrounding characters was greater for Western relative to Japanese participants and the shift in attention to the central character occurred earlier for the Japanese than American participants. Although both groups attended to the surrounding characters, Masuda et al. (2005) suggested that an Asian holistic perceptual style explained the differences.

sadness, and grief or despair experienced by an elderly man sitting at the bedside of his dying wife.

Figure 2.2 With uncertainty, one might guess the emotion expressed on the upper isolated face is boredom, sadness, or even irritation. Once the face is situated, the additional context information reveals that the expression is likely one of sorrow or grief or a blend of emotions. A single label is unimportant with situational information.

This example in Figure 2.2 illustrates the effect of situating a facial expression to influence judgments about underlying emotions. The addition of situation context activates schemas and scripts, working memory and emotions. The situational cues provide an observer with general expectations of what another individual is likely thinking or feeling, particularly when presented with blended and therefore possibly ambiguous expressions. Prior to activation of conceptual structures, evidence indicates that contextual information is integrated into the early perceptual processing of the facial expression.

In other studies, researchers showed participants facial displays embedded in background scenes and asked them to explicitly categorize facial expressions (Righart & de Gelder, 2006, 2008). In this paradigm, the faces were not naturally part of the scenes, which were independently presented in the background of the face. Congruence between the emotion expressed in the face and the scene was rapidly detected. Subjects were faster to categorize facial expressions that were congruent with the scenes behind them (Righart & de Gelder, 2006, 2008). In contrast, when the background scene elicited a different emotion than the facial expression, judgments were made more slowly and were more often incorrect. So, as with body context, social context (cartoon faces) and emotion words, emotion-eliciting scenes have been shown to affect the perceptual processing and judgments of facial expressions.

Consistent with the behavioral data, ERPs recorded during simultaneous presentations of face and scenes indicated an early influence of situational context on perception of faces that is not limited to later stage semantic associations between the face and contextual information (Righart & de Gelder, 2006; 2008). Two particular ERP components have been consistently detected in experiments: one component, the P1, has been associated with early detection of a stimulus as a face, while the other, the N170, has been linked to configural processing of the face, including facial expressions of emotion (Batty & Taylor, 2003). Importantly, an increase in N170 amplitudes was observed when expressions were presented with emotional scenes (Righart & de Gelder, 2008). Increased amplitudes of N170 were observed when fearful faces were perceived with fearful scenes relative to neutral or happy scenes. The same increased amplitude was observed in response to happy faces perceived happy relative to neutral or fearful scenes. The short latency of the N170 component indicated early discrimination of the emotional content of faces and scenes at an early stage of processing. The behavioral data showed a delay in RT to categorize happy faces with fearful scenes (Righart & de Gelder, 2006). These findings indicated affective information in the scenes was integrated with encoding of the facial expressions at an early stage of visual processing and affected the timing of explicit judgment.

The findings from the preceding studies confirm an influence of context on early perceptual processing of a facial expression, matching the perspective that perception does not rely on processing separate streams of isolated sensory cues, but rather on an integration of multiple cues from multiple sources (Zaki, 2013). Thus, the spectrum of processing required for recognition of a facial expression presented with context is expected to differ from that required for recognition of an isolated facial stimulus. This further highlights the potential weakness of many previous facial expression studies, which employed artificial, isolated facial stimuli that, because they lack context, will not adequately generalize to the facial expression recognition in one's daily experience.

The preceding review of findings about context provided compelling evidence for an effect of context on early perception and judgment about a facial expression. Since most paradigms testing for an effect of mimicry on expression recognition have not applied a situated paradigm, the inclusion of context raised questions about what role mimicry would play, if any, in situated recognition and how mimicry would interact with context.

2.3 The phenomenon of mimicry

Mimicry is the process of automatically replicating another's actions. The importance of facial mimicry in a simulation process is predicated on perception of another person's facial expression leading to automatic generation of matching facial activity in the observer and a matching emotional state. Evidence of such facial mimicry may be overt enough to be visually coded or subtle enough to require measurement of motor facial activity by electromyography (EMG) for detection. Mimicry is spontaneous, occurring without any conscious control or awareness, and can be considered externalized evidence of simulating or mirroring the state of the other person. Spontaneous motor mimicry during the observation of expressed emotion in facial displays has been well documented (Dimberg & Thunberg, 1998; Hess & Blairy, 1999; 2001; Niedenthal et al., 2010; Ponari, Conson, D'Amico, Grosso & Trojano, 2012; Sato & Yoshikawa, 2004; Stel & van Kippenberg, 2008; See also reviews by Hess et al., 1999 and Hess & Fischer, 2013), even to unconsciously perceived faces exposed for 50 ms (Dimberg, Thunberg & Elmehead, 2000) and when individuals were explicitly instructed to not generate a facial expression (Dimberg, Thunberg, & Gruendal, 2002).

There are two main perspectives on the underlying function of spontaneous mimicry. It is generally agreed that automatic behavioral mimicry (not restricted to facial mimicry) is an important social mechanism supporting social interaction, fulfilling a facilitative function to support or even enhance social interaction. One perspective is based on emotion contagion and proposes a communicative function for mimicry – to communicate sameness. The second perspective proposes a causal role for mimicry in recognition of emotional states in others.

Emotion contagion is a form of empathic responding during which an observer unknowingly "catches" the emotional state of another person and it becomes his or her own (Davis, 1984; Hatfield, Cacioppo & Rapson, 1992). One possible function of mimicry is to reflect the shared emotional state and communicate sympathy and understanding to the observed person (Bavelas, Black, Lemery, & Mullett, 1986; Chartrand & Bargh, 1999; Hatfield et al., 1992; Hess & Blairy, 2001). As a result of the emotion contagion process, the observing individual genuinely experiences a shared state with another individual and the externalized response of the shared state (e.g. in the form of a facial expression) communicates a non-verbal message of "I feel as you do" and implied understanding (Bavelas et al., 1986). In studies testing for a correlation of mimicry with empathy, and with contagion, individuals who self reported as having higher levels of empathy displayed higher levels of facial mimicry compared to individuals with low empathy (Chartrand & Bargh, 1999; Dapretto et al., 2006; Dimberg & Thunberg, 2012; Sonnby-Borgström, 2002) and higher levels of mimicry were correlated with higher levels of emotional responsivity (measured by physiological activities and self reports of feeling the emotion in response to the perceived state of another) (Balconi, Bortolotti & Gonzaga, 2011; de Sousa, McDonald, Rushby, Li, Dimoska, & James, 2011; Dimberg & Thunberg, 2012; Sato & Yoshikawa, 2004; Sonnby-Borgström, 2002). None of the studies cited above intentionally disrupted facial mimicry. The contagion perspective suggesting that overt facial mimicry communicates "sameness" and empathy has not been tested by disrupting facial mimicry.

A second perspective proposes that spontaneous mimicry is a "resonance mechanism" (Balconi, Bortolotti & Gonzaga, 2011), an early component in the motor simulation process that implements a complete embodied simulation process (Iacoboni, 2008). The function of mimicry as a resonance mechanism is thought to support recognition of the perceived emotional state of another person: Although similar to emotion contagion in that the simulated state corresponds to the perceived expression, the simulated state is typically not confused for one's own emotional state, as occurs with emotion contagion. Instead, as reported in Chapter 1, the simulated state is attributed to the other person (Goldman, 2005a, 2011), thereby facilitating the recognition process. Several studies have provided support for this hypothesis in the context of facial expressions (Adolphs, 2002; Atkinson & Adolphs, 2005; Balconi, Bortolotti, & Gonzaga, 2011; Gallese, Keysers & Rizzolatti, 2004; Goldman, 2006, 2011; Goldman & Sripada, 2005; Iacoboni, 2008; Moore, Gorodnitsky, & Pineda, 2012; Niedenthal, et al., 2001, 2010; Niedenthal, 2007; Oberman et al., 2007; Ponari et al., 2012; Rychlowska et al., 2014). For instance, when facial mimicry was intentionally blocked, recognition of facial expressions was less accurate or slower than when participants were free to mimic (Niedenthal et al., 2001; Oberman et al., 2007; Ponari et al., 2012; Rychlowska et al., 2014; Stel, Van Dijk, & Olivier, 2009). Similarly, when facial mimicry was unintentionally inhibited, for example due to Parkinson's Disease, recognition of emotion in facial expression was slower compared to controls (Livingstone, Vexer, McGarry, Lang, & Russo, 2016). However, with the exception Linvingstone et al. (2016) and Rychlowska et al. (2014), none of the studies cited had presented non-natural facial stimuli. Furthermore, none had presented expressions with scene context; as a result, little is known about the role of mimicry in facial expression recognition when scene contextual information is available. Consequently, the role of mimicry in a paradigm for situated facial expression recognition was of particular interest to the current research.

Are the two perspectives on mimicry very different?

The two perspectives on the function of mimicry – reflecting evidence of contagion and communicating sameness and understanding, and as a resonance mechanism to support simulation and recognition of another's emotional state – may not be mutually exclusive. Both are compatible in that mimicry mediates understanding and facilitates social interaction. Unlike the contagion perspective, however, in which mimicry neither facilitates nor precedes recognition of the emotional state, the resonance perspective emphasizes mimicry as a mechanism that precedes and supports recognition of an expression. According to Iacoboni, only after one internally simulates the perceived emotional state of the other person is an observer able to explicitly recognize the emotion (2008, p 112).

Effects of disrupted facial mimicry on expression recognition

If an underlying function of mimicry is to support simulation and subsequent recognition, then interference with spontaneous facial mimicry should have a negative impact on one's simulation and recognition of another person's facial expression. This negative impact could come about in a variety of ways. For example, interference in the simulation process may be due to a disruption in the afferent signals from the facial muscles to the central nervous systems (CNS) during perception (Neal & Chartrand, 2011). According to this hypothesis, feedback signals from facial expressions trigger associated autonomic and somatic responses in the brain (Adelmann & Zajonc, 1989), which in turn activate an affective state that leads to recognition. Disruption to the signals would disrupt the (shared) affect state in the self (observer) and thereby disrupt recognition.

Alternatively, the negative impact of disrupted mimicry causing interference in simulation and recognition has been explained by a disruption in the premotor and motor areas. Importantly, these areas are part of a mirror neuron system implicated in simulation and recognition. Research on mirror neurons provided evidence of common neural circuits for perception and action, suggesting that understanding another person's actions or experience and performing the same action or having the same experience, recruit common neural systems (Gallese, Keysers, Rizzolatti, 2004; Rizzolatti, Fogassi &

Gallese, 1997). The motor mirror neurons in the frontal orbital and operculum regions of the brain, particularly in Broadman Areas (BAs) 44, 45, and 6, are considered critical to a motor simulation process in expression recognition (Balconi & Bortolotti, 2012; Carr et al., 2003; Iacoboni, 2008; Jabbi & Keysers, 2008). Activity of these motor mirror neurons provides an automatic motor simulation of a facial expression based on perception. Furthermore, increased activation of the motor neurons triggered by overt motor activity caused by mimicry enhances the simulation. The idea that these motor mirror neurons simultaneously send signals to the limbic system during perception was hypothesized as the means by which the observer feels the perceived emotion (Iacoboni, 2008; Jabbi & Keysers, 2008). Activation of emotion centres (amygdala, insula, anteriror cingulated cortex), triggered by signals from the motor neurons, transmits signals to areas further downstream, including activation of autonomic and affective responses. The activation of the motor neurons are said to support resonating or mirrored emotion responding (Balconi & Bortolotti, 2012; Carr et al., 2003: Iacoboni, 2008; 2009), which in turn supports facial expression recognition (Goldman & Sripada, 2005). According to this proposed process, disruption to facial mimicry and the resulting interference with the activity of motor mirror neurons should interfere with the cascade of neural activity that supports an internal simulation of the expressed emotion and leads to recognition (Iacoboni, 2008, pp 111-113).

Past behavioral studies employing interference techniques to disrupt spontaneous mimicry have inferred a causal effect of mimicry on recognition. Following a disruption to mimicry, outcomes include impairments in categorization of static facial expressions (Oberman, Winkielman, & Ramachandran, 2007; Ponari et al., 2012; Pitcher, et al., 2008), the ability to discriminate true and false smiles (Maringer et al., 2011; Rychlowska et al. 2014) and detection of deception (Stel van Dijk, & Olivier, 2009).

Individuals who voluntarily received Botox (botulinum toxin, a fluid to intentionally paralyze activation of facial muscles) injections for cosmetic purposes showed lower rates of EMG activity and recognition accuracy for emotions compared to a control group with a full range of facial activity (Neal & Chartrand, 2011).

Intentional disruptions to mimicry also result in slower response times for activities including identification of expressions (Pitcher et al., 2008), identifications of briefly exposed static faces as being positive or negative (Stel & van Kippenberg, 2008), and detection of expression changes in computer-morphed faces (Niedenthal et al., 2001). Impaired recognition of emotions that correspond to disrupted regions of facial motor activity supported the perspective of a functional role for mimicry in expression recognition.

2.4 What was missing from the research on mimicry as a resonance mechanism to facilitate recognition?

Findings from behavioral studies support the hypothesized mechanistic role for mimicry in recognition of facial expressions. However, mimicry has not yet been tested as an important causal mechanism supporting the comprehensive process of situated facial expression recognition. A study that incorporated context in stimuli tested for, but did not find, a correlation between mimicry and recognition accuracy (Hess & Blairy, 2001). Importantly, no paradigms incorporating situational context examined a causal connection between mimicry and recognition accuracy to facial expressions. This will be a focus in the current research.

Studies reporting impaired recognition due to disruption of facial mimicry measured responses to static, black and white disembodied faces (Oberman et al., 2007; Stel & van Kippenberg, 2008), to static images of only eyes (Neal & Chartrand, 2011), and to computer-generated expressions (Niedenthal et al., 2001a, 2001b). Only recently have two studies tested for effects of disrupted mimicry on recognition of natural expressions that were dynamic (Livingstone, et al., 2016; Rychlowska et al., 2014). In those studies, participants who were unable to spontaneously mimic were slower to identify the emotional expression (Livingstone et al., 2016) or were less able to detect false rather than true smile (Rychlowska, 2014). At the time of conducting the research for this thesis there was no evidence to support a hypothesis that mimicry supports recognition of dynamic expressions, or natural, non-artificial expressions. There was only evidence pertaining to expressions that were computer generated or exaggerated posed prototypes. Such a lack of studies using natural, non-artificial stimuli raised further questions about the ecological validity of the results and therefore the strength of support for the hypothesis that the motor activity of facial mimicry supports recognition of facial expressions. Moreover, previous studies did not test the role of mimicry, and thereby simulation, on recognition of dynamic natural facial expressions that were situated; the current research aimed to address this gap by testing spontaneous mimicry as a facilitative mechanism to support recognition of situated, naturalistic facial expressions.

Is mimicry necessary in situated recognition?

Two simulation based models for recognition of facial expressions proposed by Goldman and Sripada (2005) relied heavily on facial mimicry. One model suggested a hypothetical, trial and error process of generating facial expressions until one 'matches' the perceived expression. The underlying emotional state of the matching facial expression was then attributed to the other person. The second model involved covert mimicry of an observed expression that generates a mirror state in the observer. No particular effect of a disruption to facial mimicry was offered as part of the model. Instead, alternate models were proposed for processing facial expressions under conditions when overt mimicry was not possible or required. For instance, in cases of facial paralysis due to stroke or Parkinson's Disease, recognition would be supported by a simulated response that would bypass overt motor mimicry. Instead of actual mimicry, the simulated process would involve an "as if" loop involving the somatosensory representation of what the expressed emotion would feel like, *as if* the observer had mimicked the expression (Goldman & Sripada, 2005). Additionally, in response to highly salient stimuli, information would be sent directly to the amygdala prior to cortical processing, consistent with LeDoux's proposed short circuit for emotion processing (1996). This short route model is noteworthy because it suggests that overt mimicry is not necessary for the processing of highly salient stimuli. Saliency, as a prompt for mimicry, is under debate, as discussed below.

Fundamentally, the variability in models offered by Goldman & Sripada (2005) suggests that mimicry is not always a necessary component to the expression recognition process. Niedenthal and Maringer (2009) have suggested that simulation in the form of mimicry is not necessary for simple categorization tasks, such as judging the expression as either positive or negative. However, this idea was contradicted by other findings. Ponari et al. (2012) reported better recognition accuracy on basic prototypes when individuals were free to mimic, whereas mimicry did not benefit more ambiguous expressions, such as surprise and sadness. The authors suggested surprise is often hard to identify because it could be in response to a positive or negative event and therefore is supported by either upper or lower regions of the face. Furthermore, sadness is expressed by a decrease from baseline facial muscle activity and, as a result, is difficult to recognize without context (Ponari et al., 2012). Such findings raise questions about a role for mimicry in recognition tasks, while also underlining the benefit of context for recognition of ambiguous expressions.

In contrast to Ponari et al.'s (2012) work, Niedenthal and Maringer (2009) proposed that embodied simulation in the form of facial mimicry is beneficial to recognition when detection of fine-grained differences in expressions is required. Indeed, discrimination of subtle expression differences required for recognizing true and false smiles was better when individuals were free to mimic than when mimicry was blocked (Maringer, Fischer, Krumhuber, & Niedenthal, 2011; Rychlowska, et al., 2014).

While such findings lend support to the hypothesis that mimicry is beneficial to complex recognition tasks, particularly for disambiguation of expressions by detection of fine-grained details, there is still little known about the benefit or role of mimicry in expression recognition when a natural facial expression is presented with scene context. Modulating effects of context and mimicry on emotion recognition were observed when tasks required differentiating ambiguous expressions (e.g. Aviezer et al., 2008; Rychlowski et al., 2014). Moreover, those effects occurred during early stages of perceptual processing for both context (e.g., Hassin et al., 2013; Righart & de Gelder, 2006, 2008) and overt spontaneous mimicry (790 ms after stimulus onset collected by EMG measures for healthy controls, Livingstone, et al. 2016). Thus, both mimicry and context are early parallel factors in the putative simulation processes that affect recognition of the facial expression. With early activation of both sources of input there is no evidence for either one to have a stronger influence than the other on processing the expression, especially for simple categorization. However, it has been suggested that when when facial expressions are perceived with context, particularly situational or scene context, then fine-grained discrimination of both sources of information is required (Righart & de Gelder, 2006).

2.5 Theoretical background to explain simulation processes in situated expression recognition

Three theories influenced the current work and offered insights into simulation processes that would support situated expression recognition. Although the theories represented different levels of processing, they are compatible with one another and consistent with a simulation explanation for expression recognition. These were theories for shared representations for perception and action, understanding intentions of perceived actions, and embodied cognition.

Shared representations for perception and action

The theory of perception-action coupling proposes a shared representational format for perceived events and planned actions. Based on an ideomotor framework, hypotheses for shared representations predict that perception of a particular action should simultaneously activate coding for planning and execution of the same action in the observer (Prinz 1997, 2008). Perception-action theory as the basis for models of empathy (Preston & de Waal, 2002) and has been supported by neurological evidence. Both perception and imitation of intentional actions are associated with a simple 'perception-action' neural circuit that includes the superior temporal sulcus (STS), the temporal-parietal junction (TPJ) and the inferior frontal gyrus (IFG) (Carr et al., 2003; Dapretto et al. 2006; Iacoboni, 2008, 2009). Activity in the IFG is said to trigger activity in the frontal operculum area, including activation of motor mirror neurons (Jabbi & Keysers, 2008), which as explained above, selectively fire during both perception and execution of a specific goal oriented action (Carr et al., 2003; Iacoboni, 2008, 2009; Gallese, 2003; Gallese, Keysers, & Rizzolatti, 2004; Rizzolatti & Craighero, 2004; Rizzolatti, Fadiga, Fogassi & Gallese, 1996). The functional overlap in firing of these neurons is consistent with the representational correspondence predicted by the ideomotor model. Thus, an important proposed function of the mirror neurons is to support the observer's understanding of the intention behind specific goal oriented actions in others (Rizzolatti et al., 1996; Gallese, 2003; Gallese, Keysers & Rizzolatti, 2004; Iacoboni, 2008). With inclusion of areas associated with emotion processing, such as the amygdala, insula and anterior cingulate cortex (ACC), the perception-action circuit is also thought to support recognition of the emotional meaning of actions, including facial expressions (Carr et al, 2003; Iacoboni, 2008) and empathic responding (Preston & de Waal, 2002). Iacoboni suggested that signals from IFG mirror neurons that connect to the limbic system via the insula would support perception of emotion within the facial expressions of others, resulting in the observer feeling the same emotion (Iacoboni, 2008, pg 111).

In addition to the example included in Chapter 1 of common coding for experiencing and perceiving disgust (Jabbi & Keysers, 2008; Wicker et al., 2003) evidence of perception-action coupling in emotion processing was found in reports of common neural activity for experiencing non-affective pain and perceiving another person experience non-affective pain (Singer et al., 2005). Activation of the amygdala has frequently been associated with both perceiving and experiencing high levels of arousal to positive as well as negative stimuli (Adolphs, 2002; Kennedy & Adolphs, 2012). These findings support a slightly expanded neural circuit for perception-action that includes emotion-processing areas (Carr et al., 2003).

A cognitive mechanism for integrating representations of actions

The second influential theory for this research was proposed by Barresi and Moore (1996, 2008) and features a cognitive mechanism that integrates information about a perceived intentional action and execution of the same action. Perception activates in the observer a schema for that action that supports matching of the first and third person representations of the action. Similar to the perception-action theory outlined above, the first-person representation of the perceived action provides the observer with information as though the observer were generating the action him or herself. Integration of the two sources of information into a single representation is thought to support recognition of the intention behind the perceived action. This cognition based theory is consistent with simulation theory and compatible with the theories for shared representations for perception and action, but puts greater emphasis on the intention of the action with respect to a given situation that is usually current but sometimes imagined or remembered. The temporal and conceptual aspects of Barresi and Moore's theory make it compatible with the following embodied cognition model for understanding the meaning of perceived actions.

Embodied cognition

The theory of embodied-cognition is particularly helpful in explaining the underlying processes for facial expression recognition and, in particular, how the observer's response to contextual information is incorporated into his or her recognition of a facial expression. Proponents of embodied cognition suggest that cognition is a sensory experience supported by multiple systems throughout the entire body (Winkielman et al., 2008). For example, the early stages of perception of a smile would involve more than just the visual system; perception activates conceptual information based on memory or learning about the meaning of the symbol of the smiling face as well as motor systems associated with production of a smile, which in turn activate affective and autonomic systems associated with the positive stimulus. Activation of a multiple system response during perception suggests that perception and recognition of a facial expression comprise an embodied process associated with common systems generating the expression.

Embodied cognition is a particularly helpful theory for explaining underlying processes that support situated expression recognition, as the theory provides an account for the incorporation of contextual information into the meaning of perceived events and stored in memory. It has been suggested that motor, somatosensory, and affective states triggered during an original experience with an emotion-eliciting stimulus are captured, stored and reactivated again during later observation or recollection of the same or similar stimulus (Barsalou, 1999; Niedenthal, 2007; Niedenthal, Barsalou, Winkielman, Krauth-Gruer & Ric, 2005; Niedenthal & Maringer, 2009; Winkielman, Niedenthal & Oberman, 2008). For example, an encounter with an assailant would typically activate a fear response, with neural coding that involves activation of multiple systems (e.g., perceptual, somatosensory, interoceptive, and motor resources) that contribute to the cognitive processing of and behavioral response to the fearful event. Later perception of images of an assault or remembering the same assault could re-activate, at least partially, the original neural coding or pattern of sensorimotor and affective states that occurred during the initial assault experience (Niedenthal, 2007). Reactivation of the coding generates a simulated experience, one that partially mirrors the original experience. Because such events occur in context rather in a vacuum, multiple modes of information about a given situation are incorporated into the processing of the initial experience (e.g. time of day, scent, surrounding sounds, visual details of the location, etc). Thus, this reactivation could be caused by information not central to but nonetheless related to the initial experience.

Reactivation of associated states could be triggered in response to corresponding contextual stimuli perceived with a facial expression. Reactivation of the original coding may be partial, but what is reactivated is thought to depend on what information is currently relevant and selectively attended to by the observer (Barsalou, 1999; Preston &

de Waal, 2002). The later, reactivated experience is referred to as embodied because although the experience is not overtly reproduced, parts of the prior experience are covertly reproduced by reactivation of the original multimodal coding, as if the observer were in the very situation (Winkielman, Niedenthal & Oberman, 2008).

Given that, according to embodied-cognition theories, conceptual processing recruits multimodal systems (Niedenthal, 2007; Niedenthal, Barsalou, Winkielman, Krauth-Gruer & Ric, 2005; Niedenthal & Maringer, 2009; Winkielman, Niedenthal & Oberman, 2008), it can be argued that conceptual knowledge (i.e. knowledge supported by learning, inferences, or imagination) can trigger visceral and physiological activity in response to an emotion-eliciting stimulus without overt prior experience. For example, an image of an individual falling off a cliff is a fear-eliciting stimulus that without past explicit experience of falling off a cliff could activate in the observer a multimodal representation that stimulates a change in heart rate, skin conductance, and possible motor response such as “pulling one’s self back to prevent the fall”. One can imagine the experience of the individual in the image by projecting oneself into that situation, and conceptual knowledge enables one to imagine the consequences.

Embodiment of conceptual knowledge in expression recognition

Of particular relevance to this thesis, the recruitment of conceptual knowledge, without explicit experience, can activate an embodied, multimodal representation of a perceived event or object. An embodied response to context perceived with a facial expression is expected to play an important role in the underlying processes for recognizing the emotional state of the observed individual. In a live situation, the observer’s response to context is expected to interact with their response to behavioral cues perceived in another person’s facial expression. Furthermore, in a live situation in which both the observer and observed individual experience the same context (e.g. sudden activation of a fire alarm), the observer's embodied response to shared context is said to resonate with the state of the other person, which supports comprehension of the emotional state of that other person (Gallese, 2003).

Implicit recognition of the facial expression was thought to be an automatic, bottom-up response, likely supported by neural activity in regions associated with shared representations, such as the mirror neuron system (MNS) (Gallese, 2003, 2007; Iacoboni, 2006; Spunt & Lieberman, 2011; Shamay-Tsoory, 2011). Another circuit associated with shared representations is the mentalizing system. Activities like judgments about an expression as a likely response to situation involves more elaborate and complex processing, and are associated with the mentalizing system (Shamay-Tsoory, 2011; Spunt & Lieberman, 2011; 2012; Van Overwalle & Baetens, 2009).

Spunt and Lieberman (2013) suggest that explicit judgments about an expression are supported by the mentalizing system (MS), however, activation of this system is also associated with preconscious, inferential processing (Malle, 2005). Much of the MS is associated with formation of inferences about another individual's mental state, and the ability to imagine that person's experience or perspective. Moreover, inferences about beliefs and emotional states are not limited to referencing another individual, as they are also formed in reference to the state of the self. Important to the processes under scrutiny for situated recognition, the MS is associated with production of representations of self and other (Olsson & Ochsner, 2007; Shamay-Tsoory, 2011, Spunt & Lieberman, 2011; Van Overwalle & Baetens, 2009). Although this system is not readily associated with embodiment, it is suggestive of processes that support the reflective component of one's response to a context scene.

Fundamentally, accounts of embodied cognition suggests that comprehension of an observed person's emotional state depends not just visual cues perceived in the person's facial expression, but also the context in which they are observed, and moreover, also on the observer's response to the same context (Winkielman, Niedenthal & Oberman, 2008). The principal of an observer's embodied response to a shared context informing their understanding of the meaning of another person's facial expression is applied to a situated expression recognition task: An observer's response to a context scene presented with a facial expression is expected to inform judgments about the meaning of the expression.

While some simulation theorists refer to the simulated state that resonates with another individual as a current state without overt emphasis on reactivation of any past experience (e.g. Gallese, 2003), other simulation theorists, particularly those influenced by embodied cognition, recognize the role of partial re-activation of a past experience or conceptual knowledge involving excitation of multiple systems (e.g. Niedenthal, 2007; Niedenthal, Barsalou, Winkielman, Krauth-Gruer & Ric, 2005; Niedenthal & Maringer, 2009; Winkielman, Niedenthal & Oberman, 2008). Consistent with this embodied-simulation perspective that incorporates past experience, theories of perception-action coupling (Carr et al., 2003; Prinz 1997, 2008) and cognitive integration of the meaning of perceived and executed actions (Baressi & Moore, 1996, 2008) suggest an implicit criterion of knowledge and/or shared past experience with the given actions, state, or situation of the observed person. The implication is that the degree of knowledge or experience influences the strength of one's representation and comprehension of the meaning of the perceived or imagined state of the observed person. Fundamentally, what may vary in the observer's embodied response is the degree to which re-activation of the original neural pattern of the sensorimotor or affective state occurs. For instance, a greater degree of salience, previous experience, or familiarity with the perceived stimulus is predicted to result in greater reactivation of the original neural coding and a more robust representation of the perceived event or object of one's attention (Preston & de Waal, 2002).

2.6 Summary of the theories and their application in a situated expression recognition paradigm

In summary, perception-action coupling offers an account of common coding for perception and imitation of facial expressions, with particular emphasis on the role of motor mirror neurons. Barresi and Moore's theory (1998, 2006) offers a model of cognitive mechanism to support integration of visual cues that are directly available to the observer with his or her own matched, internalized responses to a shared context, be it current or conceptual. The embodied cognition model similarly suggests that recognition is supported by partial reactivation of a past experience or activation of a multimodal representation of conceptual knowledge that matches current perceptual information. These three theoretical models are consistent with simulationists' interpretation of emotion recognition in facial expression wherein perception of visual cues automatically activates a mirrored state in the observer (Goldman & Sripada, 2005).

Simulation theorists have suggested that, through partial reactivation of neural coding of an expressed emotion, the observer simulates the emotional state of the other person and assigns the simulated state not to him or herself, but to the other individual (Goldman, 2005; 2011; Goldman & Sripada, 2005), which in turn supports the observer's recognition of the state of the other person. In contrast, the perception-action, cognitive and embodied-cognition models support a mirroring process for recognition of a perceived facial expression as a response to situational context that is shared with the observer, as described above. In this thesis, aspects of perception-action coupling, cognitive integration, and embodied cognition theories were integrated into a single model for situated facial expression recognition.

An actual example of live situated expression recognition involves rapid recognition of the emotional state of another individual by integrating one's own response to a shared context with their representation of the other individual's facial expression that is in response to same context. Simply put, understanding the meaning of the expression is partially derived by unconsciously referencing one's own response to the same event. As a simple example, two people are watching a dog chase it's tail. Person A experiences pleasure by the dog's actions. Person A perceives Person B's facial expression. A's representation of B's facial expression matches A's internal response (embodied state of pleasure) to the dog's behavior. Person A recognizes B's emotional state as pleasure.[4]

Situated expression recognition is tested for in this thesis by presenting pairings of context scene and a facial expression. The participant must recognize the expression as the response (or not) to the scene. As a hypothetical example, a picture of a dog chasing his tail is the context scene presented with an image of a facial expression. Specifically, the task is to determine if the expression is in response to the scene of the dog chasing its tail. It was proposed that an understanding of the meaning of the perceived facial expression is supported by one's simulation of the expression. The embodied simulation of the expression results in a representation of the expression that is recognized as a match (or not) to one's own internalized, embodied response to the context scene. The processes underlying this matching are thought to occur rapidly and below the threshold of consciousness (Barresi & Moore, 1996, 2008).

How recognition of the match between a perceived facial impression and the embodied recognition occurs is of particular interest to the current research. If the

[4] This example of situated expression recognition is intentionally simple and the proposed underlying processes are explained in greater detail in Chapter 3.

representation of the facial expression is generated by a mirrored state in the observer (Gallese, 2003; Goldman & Sripada, 2005; Iacoboni, 2008; Preston & de Waal, 2002), then facial mimicry may indeed be an important component in that mirroring. Is this really the case? If facial mimicry isn't possible due to injury or illness, is the mirrored state sufficient without mimicry? Is mirroring by mimicry a more effective process for recognition than formation of inferences associated with the mentalizing system?

There is a dearth of information in the literature explaining the underlying processes for situated facial expression recognition and, in particular, the role of mimicry in those processes. The testing of mimicry was been conducted with conceptual context (emotion words, Halberstadt et al., 2009) and social context in the form of gender, Hess & Fischer, 2014) and social dominance (Carr, Winkielman & Oveis, 2014) and more recently with beliefs as context (Maringer, et al., 2011). None of these studies tested mimicry with scene context in the paradigm, nor was recognition tested to natural expressions. Only recently had Rychlowska et al., (2014) tested for an interaction of mimicry and beliefs on recognition of natural expressions. Moreover, although context had been incorporated into the task, there was no mechanistic explanation for how context would impact recognition.

Thus, as previously stated, the research for this thesis was motivated to test the theory that a simulated state supports situated natural facial expression recognition. In particular, this research aims to address the question of how critical spontaneous facial mimicry is to a situated recognition process.

In an effort to better comprehend and explain a situated recognition process, another issue is whether one individual's perception and interpretation of another person's facial expression is influenced by his/her own embodied response to a shared context. This possibility suggests that an embodied response to a given context primes perception of the facial expression of another person observing that context, thereby influencing judgment about the perceived expression of the other person.

Although past research has measured the effects of mimicry and context on recognition, mimicry and scene context have not been tested together. For the purposes of this thesis, a simulated response to a perceived facial expression and an embodied response to a shared context were proposed as important factors in situated facial expression recognition.

CHAPTER 3: THE CURRENT WORK

The objective of this thesis was to investigate the underlying processes for situated expression recognition. In doing so an explanation for expression recognition offered by simulation theory was tested. To that end, the research was built on the premise that perception of another person's facial expression automatically activates a corresponding, simulated emotional state in the observer that is automatically attributed to the other person (Gallese & Goldman 1998; Gallese, 2003, 2007; Goldman 2005).

The preceding chapters outline substantial research into evidence for automatic mimicry and the role of mimicry in expression recognition. Facial mimicry has been identified as a resonating mechanism (Balconi & Bortolotti, 2012) underlying recognition with an important function to strengthen levels of motor activity triggered by perception of the expression, and thereby enhance the continuing simulation of the perceived expression.[5] As reported, the findings were inconsistent and testing stimuli were generally static or artificially generated images. In this thesis, it was questioned if mimicry underlies recognition of natural expressions that are situated.

In situated expression recognition the observer's response to a shared situational context is identified as another important component in the recognition process. Research reviewed in Chapter 2, section 2.2, provided significant evidence that context is detected early in perceptual processing of a facial expression (Hassin, Aviezer, & Bentin 2013; Righart & de Gelder, 2006; 2008), and the effect of perceived context was

---

[5] It is important to clarify that the simulated emotional state is reportedly achieved by activation of multiple systems in addition to the motor system. Thus, the simulation is considered an embodied process. Many theorists refer to the process as embodied simulation but for the sake of clarity in this thesis, the response to a perceived facial expression is referred to as simulation and the response to the context scene is referred to as the embodied response.

observed on early encoding and later recall of a facial stimulus (Halberstadt, et al., 2009). These preceding arguments for an effect of context influenced questioning if the observer's response to context underlies recognition of an expression.

Hence, two hypotheses were tested throughout the research. The first was a hypothesis of automatic facial mimicry as an underlying mechanism in situation expression recognition. This hypothesis was tested by disrupting mimicry. If mimicry is an important component in the simulation process, a disruption to mimicry should impair recognition.

The second hypothesis was of the observer's response to context as an underlying component in recognition process. In the current recognition task, facial expressions were judged as the true or false response to a context scene. Measuring judgments of an expression as the true or false response when the expression and scene elicited the same or similar emotion tested this hypothesis of context. Under these conditions, an influence of context would result in significantly higher accuracy rates to judge the expression as true. Details of the recognition task are outlined below. Additional details about the hypotheses and predictions are addressed in the final section of this chapter.

3.1 The paradigm for situated expression recognition

Motivated by the need for increased ecological validity in expression recognition research, the facial stimuli used in the research for this thesis were not static prototypes but rather colour videos of naturalistic, spontaneous facial responses that ranged in intensity and expressivity. The facial stimuli were not disembodied, instead including the full face, head and shoulders.

Natural expressions are anticipated to be ambiguous relative to prototypes used in most studies, because of their subtlety and blended emotions. The videos were from a bank of expressions composed of four models' previously recorded expressions in response to standardized scenes of emotional content. The emotional content was drawn from the International Affective Pictures Systems (IAPS) collection, some of which had been rated as eliciting more than one emotion (e.g., anger and disgust) (Lang et al., 2005).

Social interaction partners are motivated to attribute another person's expressed emotion to a causal factor (Hieder, 1958) and a context scene helps to disambiguate an expression by providing a situational meaning for the expression. In this research, context scenes were presented with the models' facial expressions in order to better represent a social interaction.

Figure 3.1 On a single trial one scene is presented with one model's facial expression that is or is not the actual response to that scene.

Previous studies testing for an effect of context by embedding a human face on cut-out images of human bodies (refer to examples by Aviezer, in Ch 2.2) were criticized as unnatural (Matsumoto, 2010). Consequently, providing images of situational context that might represent the natural world was preferred. The natural scenes to which the models responded were included in the paradigm as situational context and provided information about what caused the model's emotional response.

Thus, a paradigm was designed for testing situated facial expression recognition. Over multiple trials, a static image representing situational context was presented simultaneously with one dynamic (video) facial expression. The observer's task was to judge whether or not the facial expression was the model's "true" or "false" response to the context scene. Facial expressions were not random matches. Fifty percent of trials paired a context scene with a model's "true" response to that scene (the model's actual response to the scene) and 50% of trials paired a context scene with a model's "false" response (the actual response to one other scene).

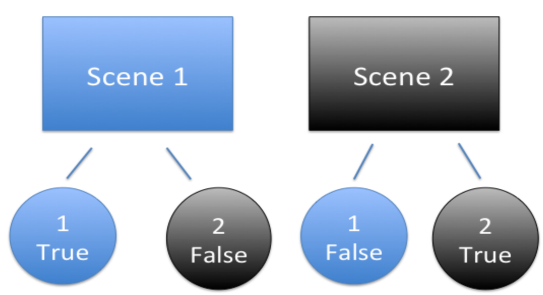

Figure 3.2 On 50 % of trials a context scene is presented with a true response, and on 50% of trials a scene if presented with the false response. Stimuli are organized such that context scenes were paired, which meant that each expression was presented with 2 scenes. The true response to scene 1 is the false response to scene 2.

Recognition on the part of the observer was measured as the ability to accurately judge the expression as being true or false with respect to the context scene, responding to the question “Was this person looking at that image?” with a forced choice of “Yes” or “No”.

Expression labeling or categorization was not part of the paradigm because the research for this thesis was primarily concerned with the attribution of mental states that support expression recognition. Labeling involves a verbal component that is not critical to the attribution process. Furthermore, preceding arguments in Chapter 1 regarding an increased risk of experimental artifacts caused by labeling supported the decision to avoid labeling in the task.

In this thesis, a multi-stage task is proposed for testing situated expression recognition. In the task, the observer’s simulation of a perceived expression results in a representation that is then implicitly recognized as a match (or not) to the observer’s internalized, embodied response to the simultaneously presented context scene. Figure 3.3 shows a two-fold process involved with judging whether the model’s expression is a true or false response to the context scene.

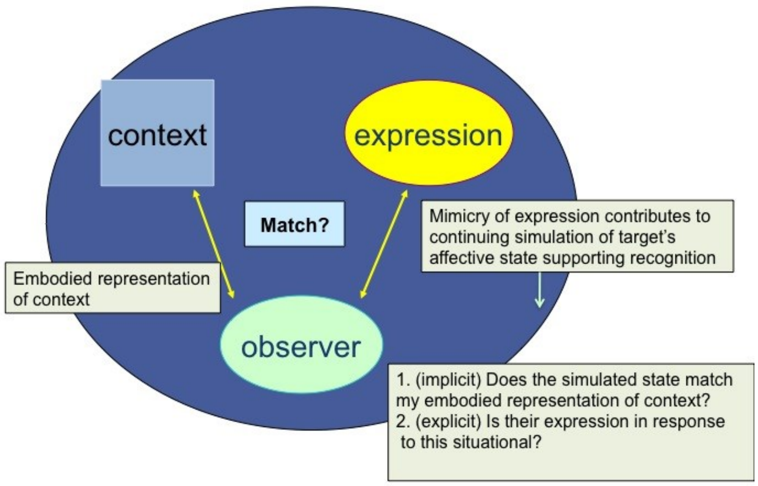

Figure 3.3. A simple schematic of the two-fold process in the task, detection of recognition followed by attribution

Using very broad terms to describe the task, the process could be described as recognition and attribution. However, explaining the finer details of the situated recognition task reveals multiple stages and parallel processes that occur between detection and recognition, and then between recognition and judgment of the expression as a response to the accompanying scene. The suggested underlying processes for this task follow the proposed sequence of recognition as outlined in Chapter 2.3. That sequence was defined according to the principles of shared representation for perception and action, embodied cognition and cognitive integration of the observer's representation of the facial expression with his or her internal response to the context scene. The sequence is illustrated in Figure 3.4 and is described in detail in the paragraphs to follow.

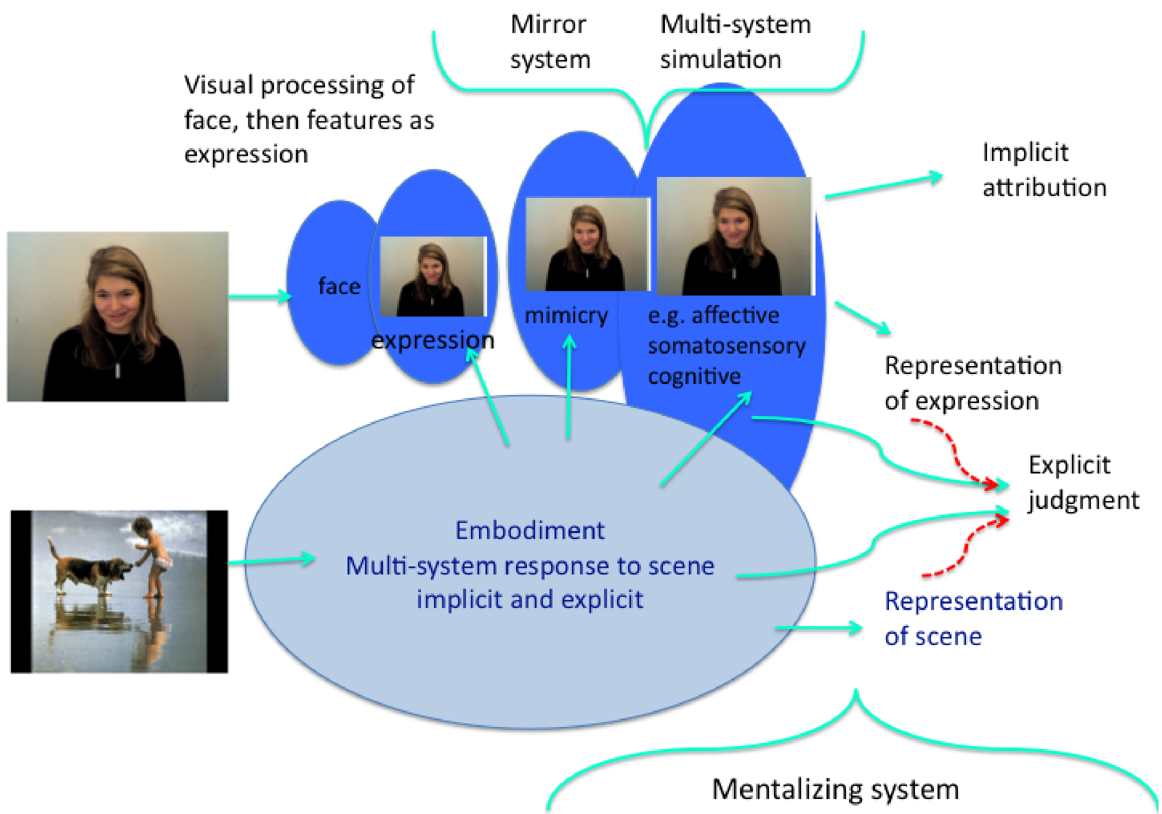

Figure 3.4. A proposed schematic for the sequence of processing a facial expression and context scene that supports judgments about the expression as the true or false response to the context scene. The two images on the left represent the facial expression and context scene that are presented in the situated recognition paradigm.

An early part of the judgment process involves rapid detection and implicit recognition of the facial stimulus as a face, with a particular configuration of features that signal an expression. Perception of the facial expression triggers activation of the mirror neuron system and automatic facial mimicry. Facial mimicry increases the motor signal to additional systems to enhance ongoing simulation of the perceived expression. The simulation process generates a corresponding emotional state in the observer, which is automatically attributed to the model.

Concurrent is the participant's processing of the context scene. Activation of common multiple systems also supports the embodied processing of the scene. Autonomic responses, relevant schemas and memories are activated, and feelings arising from the embodied response reach awareness.

At a subthreshold level, correspondence between the represented emotional state of the model and the observer's response to the context scene is implicitly recognized as a match. An explicit judgment about the expression as a response to the context must occur. The preconscious recognition of the match (or not) supports the explicit judgment to follow.

In addition to the mirror neuron system supporting the simulation of the expression, another system proposed as likely active during the final stages of recognition is the mentalizing system (MS). Recall that the MS is associated with formation of inferences about emotions and beliefs and the ability to imagine another person's experience or perspective. Importantly, the inferences formed by the MS are about the mental states of another person as well as the self. (Olsson & Ochsner, 2007; Shamay-Tsoory, 2011, Spunt & Lieberman, 2011). The MS is mentioned here with the intention of explaining the fine details in the final stage of the complicated recognition process. It is proposed as likely active during attribution of the simulated emotional state to the model, the participant's awareness of their own response to the context scene, and formation and integration of the their representations of the model's expression and their own response to the context scene.

These implicit assessments are followed by an explicit judgment of the expression as a response to the context scene. Hence, the mentalizing, both explicit and implicit, associated with the MS suggests a role for this system in the task. The task anticipates preconscious mentalizing by a participant about his or her own response to context scene as well as the model's response. Controlled assessment follows preconscious assessment, resulting in an explicit judgment about the expression as a response to the context scene (Shamay-Tsoory, 2011). The inclusion of this system provides the fine details for explaining how the simulated response to the expression is integrated with the embodied response to the context scene, and supports the final stage of the participant's recognition of the represented emotional state of the model as a match, or not, to his or her own.

To describe the recognition process in an externalized, ecological way, it was expected that the observer would experience his or her own response to the context scene and search for evidence in the expression of the model for an underlying emotional state that matches the observer's response to the scene. A matching state suggests that the model had indeed been looking at the same context. A perceived shared state would prompt a judgment of the model's response as being a 'true' response to the image.

3.2 Predictions

On the strength of preceding evidence in Chapters 1 and 2, it was possible to form certain predictions regarding the outcomes of the current study. Hypotheses that are particular to each of the three studies are presented in each chapter. Two main hypotheses continue through the research.

The first continuing hypothesis is of the participant's response to context underlying recognition of the expression. Based on existing evidence of the modulating effects of context on early perceptual processing of facial expressions (e.g., differing amplitudes of N170 with different face context combinations indicated an early impact of context on face processing, Righart & de Gelder, 2006, encoding, and later recall of expressions, Halberstadt, et al. 2009) processing the context scene is expected to influence judgments about the expression. The participant's embodied processing of the scene is expected to occur in parallel, if not overlap, with processing of the facial expression. The observer's embodied response to the context scene is expected to influence implicit as well as explicit judgments. The explicit judgments are measurable.

Discrimination of the model's response as true or false is difficult when the context scene represents either the same or a similar emotion. In such a case, processing the expression and context scene are likely to interact, and influence the final explicit judgment.   An influence of context on perception of the expression would manifest as a tendency by participants' to judge the response as "true" – as a match to his or her own response to the context. An index of such an influence was measured as higher levels of accuracy for judging responses as "true" than for judging responses as "false" on trials with expressions and context scenes of the same valence. In the current work, valence refers to the positive or negative quality of arousal caused by the stimulus. Evidence of such an influence in the data will be confirmed with additional criterion analysis.

The second hypothesis is of mimicry underlying expression recognition. As reviewed in Chapters 1 and 2, mimicry has been proposed as a possible mechanism to support the recognition process, by means of either activating a resonating mechanism (Balcon & Bartolotti 2012) and generating increased motor activity that enhances simulation (Iacoboni, 2008; 2009). As reviewed, past studies have reported a beneficial effect of facial mimicry on expression recognition by disrupting mimicry, but these studies used non-natural facial stimuli that were not presented with scene context. Consistent with these prior testing procedures, the hypothesis of a role for mimicry in recognition will be tested by disrupting mimicry. If mimicry does support recognition of naturalistic expressions in a situated paradigm, then recognition of the model's response as either "true" or "false" will be more accurate when participants are free to mimic than when they are not. On the other hand, if mimicry is not beneficial to recognition, there should be no difference in accuracy between disrupted and free to mimic conditions, and the hypothesis that mimicry underlies recognition of naturalistic expressions in a situated recognition task would not be supported.

Predictions about an interaction of mimicry and context were less certain. Maringer et al. (2011), and more recently Rychlowska et al. (2014), reported an effect of disrupted mimicry on recognition of smiles as true or false. Results supported their hypothesis that mimicry was beneficial in detecting cues that enabled recognition of smiles as false. When mimicry was disrupted, participants relied on their beliefs about the incentives of behaviour to guide their judgments of smiles as true or false. However, the facial expressions presented by Maringer et al. (2011) were avatars. While facial expressions used by Rychloskwa et al. (2014) were natural, they did not include scene context in the task. With their findings in mind, in the current study, it is expected that an influence of context will be most evident when mimicry is disrupted. If mimicry does underlie recognition, then based on the findings of Maringer et al. (2011) and Rychloskwa et al. (2014), it is expected that a disruption of mimicry will decrease detection of fine-grained cues in the perceived expression. Without mimicry, participants will instead rely on their own response to context to guide their judgment about the meaning of the model's facial expression, resulting in an even stronger to judge the model's response as true during conditions of disrupted mimicry.

3.3 Power Analysis

Although an *a priori* effect size was not estimated for the first two studies, the goal was to collect data from a sample size that matched the *N* of relevant, comparable studies with a mixed design consisting of two groups. Several studies that had reported effects of disrupted mimicry on recognition and applied a mixed subject design with two groups, also reported total *N*s of 97 (Niedenthal et al, 2001), and 60, (Hess & Blairy, 2008). Two later but even more comparable studies with two groups had also reported total *Ns* of 64 (Maringer et al., 2011) and 66 (Rychlowska et al, 2014). Hence, with the comparable sample sizes for Experiments 1 and 2 there was confidence of having adequate power and avoiding a Type I error. The design for the third experiment, reviewed in Chapter 6, resulted in four groups. For this study, a sample size of 96 provided adequate power based on *a priori* calculations for a large estimated effect size of .80, with an alpha of .05 (Cohen, 1992).

Chapter 4: Experiment 1

A functional role for facial mimicry in facial expression recognition has been demonstrated several times by experiments interfering with the activity of facial muscles. The results of such studies, demonstrating higher accuracy when participants were free to mimic relative to when mimicry was disrupted, supported the hypothesis of a facilitative effect of mimicry. However, it is difficult to interpret these studies as a whole because, although their conclusions are similar, their methodologies vary substantially. Furthermore, none of these studies included context in their experimental paradigms. Studies that tested the effect of mimicry on expression recognition were influenced by earlier research on emotion priming. Those studies are outlined below in order to provide an overview of the existing literature and to establish a rationale for the manipulations used to disrupt facial mimicry in the first experiment of the current work.

Early studies tested activated facial muscles to test for an effect of induced emotional state on judgments of perceived events. The early studies were based on principals of facial feedback. The manipulations were then applied to research on expression recognition. Working from principles of embodied cognition and simulation theory, several studies disrupted facial activity and tested for an impairment effect on recognition. These procedures for inhibiting or interfering with facial mimicry have been replicated for research on mimicry and embodied simulation. The rationale and relevant findings for three main studies are reviewed and compared below. Their findings and procedures were influential on the research and methodology for this first Experiment. A review of their research is followed by the details for Experiment 1.

4.1.i. Early tests of a manipulation to prevent facial muscle movement during facial expression processing

Relevant findings about the effects of facial activity, like smiling, influencing social judgments had an impact on paradigms for facial expression recognition. Previous investigations have shown that the emotion experienced by an observer at the time of encoding a stimulus, such as a face, can alter perception and encoding of that stimulus (Ambady & Gray, 2002; Barrett, Mesquita, & Gendron, 2011; Moody, McIntosh, Mann & Weisser, 2007; Halberstadt & Niedenthal, 20001; Halberstadt et al., 2009; Niedenthal et al., 2000; 2001; Wild, Erb & Bartels, 2001). Similarly, producing a facial expression has been shown to change an observer's emotional state, as evidenced by affective self-reports (Dimberg & Thunberg, 2012; Strack, Martin & Stepper, 1988), ratings of affective stimuli (Laird, 1974), and changes in autonomic conditions such as heart rate and skin conductance (Ekman, Levenson, & Friesen, 1983; Kraft & Pressman, 2012). For instance, intentional activation of the zygomaticus major muscles, which are typically activated during smiling, has been associated with more positive ratings of cartoons (Laird, 1974; Strack et al, 1988), self-reports of a positive state in response to a positive expression (Dimberg & Thunberg, 2012), and lower levels of a stress response (Kraft & Pressman, 2012). This phenomenon has been explained by afferent feedback to the brain generated by facial activity (Adelmann & Zajonc, 1989: Kraft & Pressman, 2012) (previously discussed in Chapter 1). This 'facial feedback' hypothesis has been tested by manipulating facial muscle activity and recording individuals' affective responses to various stimuli (Laird, 1974; Strack, Martin & Stepper, 1988).

In one such study, Strack, Martin and Stepper (1988) tested the hypothesis that facial muscle contractions associated with smiling influence affective responses by having individuals rate the "funniness" of cartoons while holding the *end* of a pen in their mouth. Participants held the pen either between their teeth, which was intended to facilitate smiling, or between their lips, which was intended to inhibit smiling. The results revealed that those who held the pen between their teeth, causing activation of the zygomaticus major, rated cartoons as funnier than those who held the pen between their lips, which inhibited activity of the same muscle. Furthermore, those who held the pen between their teeth also reported "feeling" more positive than those who held the pen between their lips. Strack and colleagues (1988) interpreted their results as support for the facial feedback hypothesis

Also based on the principle of facial feedback, a series of experiments tested an effect of emotion congruence on the speed at which observers detected changes in facial expressions (Niedenthal, Halberstadt, Margolin, & Innes-Ker, 2000: Niedenthal, Brauer, Halberstadt, & Innes-Ker, 2001). Participants first experienced an emotion induction procedure intended to produce a happy or sad emotional state; they then watched movies of morphed facial expressions that changed from happy to sad or from sad to happy (Niedenthal, et al., 2001). Participants were tasked with marking the exact frame at which an expression changed (e.g. from happy to sad, participants had to select the frame representing the offset of happiness). An emotion congruence effect was observed, such that participants in the induced happy state detected the offset of happiness earlier than those in the sad state, and conversely, participants in the sad state detected the offset of sadness earlier than those in the happy state. This was consistent with Strack et al.'s (1988) interpretation regarding the facial feedback hypothesis, as a primed emotion-congruent state influenced perception and judgment of the stimulus.

Niedenthal and colleagues (2001) claimed that faster expression detection in the congruent condition was likely due to facial mimicry and differences in facial feedback.

Mimicry may have facilitated the detection of change in the initial expression because the onset of a new, incongruent expression produced a noticeable shift in the perceiver's facial activity, and feedback caused by that shift influenced the ability to detect changes in the expression (Niedenthal et al., 2001).

Niedenthal et al. (2001) also tested the hypothesis that mimicry influences perception of a facial expression by applying a manipulation intended to prevent mimicry. Participants completed the same task as above, identifying the frame at which the facial expression first changed; however, in this experiment, half the participants held a pen "sideways in their mouth, lightly holding it with their lips and teeth" during the task. This manipulation, modified from the neutral condition of Strack et al. (1988), had been previously shown to prevent mimicry that would simulate a smile but does not prevent mimicry associated with the upper areas of the face. Thus, some aspects of the facial activity could be mimicked despite this disruption. Participants who were free to mimic detected change in expression earlier than participants who could not mimic. Once again, the researchers suggested that the faster detection was due to mimicry, with those free to mimic experiencing valuable feedback due to changes in facial activity (Niedenthal et al., 2001).

Strack and colleagues (1988) demonstrated that manipulation of facial activity modulates affective judgments. Enhanced production of a smile generated feedback signals that triggered a positive state in the participant and influenced positive ratings of a stimulus. The results supported the facial feedback hypothesis in that enhanced facial activity generated feedback signals that primed an affective state of the participant and influenced their judgment of a stimulus.

In one study, Niedenthal et al. (2000) also found a priming effect of participant's affective state on perception of facial expressions. In their follow-up study the researchers reported that feedback signals generated by spontaneous mimicry triggered an affective state in the participant that was congruent with the displayed expression. As the expression changed, mimicry generated updated feedback signals that supported rapid detection of the change (Niedenthal et al., 2001).

Results of both studies supported the facial feedback hypothesis. However, the methodology of the study testing mimicry raised questions about the validity of the results. The manipulation with a pen was well suited to interfere with a detection of change from happiness. Spontaneous mimicry of a happy expression typically activates the zygomaticus muscle, whereas mimicry of a sad expression does not. Thus, any change from happy to sad or sad to happy would be rapidly detected by the slightest change in zygomaticus activity. It was not clear that any additional downstream activity, such as sharing a congruent affective state, was actually required for their recognition task, especially given their procedure of a frame-by-frame analysis and unlimited time for the task. It was not entirely convincing that a congruent affective state in the participant was necessary for detecting change in the expressions, as suggested by the authors. However, at the very least, disruption to facial motor activity did seem to delay detection of change in expressions of opposite valence emotions.

Importantly, Niedenthal et al. (2001) identified mimicry as a mechanism to influence the perception of facial expressions. It was unclear, however, if mimicry supports recognition of a change in expression when the emotions are very similar. Thus, their evidence for the role of mimicry in recognition of expression was neither convincing nor absolute. Nonetheless, their study became a landmark that inspired subsequent research testing for a functional role of mimicry by disruption to facial activity.

4.1.ii. Additional research on the role of embodied simulation in emotion recognition

In direct response to the research of Niedenthal et al. (2001), Oberman, Winkielman and Ramachandran (2007) generated a more focused hypothesis regarding the role of embodied simulation in the recognition of facial expressions. They proposed that, because different emotions vary in the degree of expressivity in the face, blocking facial mimicry should *differentially impair recognition of different expressions*. For instance, the expression of happiness is associated with many areas of activity in the face (Ekman, 2004) and should therefore be externally simulated, drawing on the perceiver's own facial musculature, to a greater degree than sadness, which is associated with activity in the face that is more subtle. Consequently, the simulation of sadness may be a more internalized experience, relying relatively little upon motor activity of the perceiver's face for understanding (Oberman et al., 2007).

Oberman and colleagues (2007) predicted that interfering with a simulation process by generating *irrelevant* facial motor activity ("muscular noise") should be most detrimental to recognition of expressions that involve substantial facial activity, such as happiness, and have little impact on recognition of expressions that involve less facial activity, such as sadness. Initially, they measured the activity of four facial muscles during expressions of happiness, sadness, fear and disgust.[6] Electromyographic (EMG) measures confirmed that, of those four expressions, happiness generated the most facial activity and sadness generated the least activity. Thus, it was predicted that interference

---

[6] The four muscles measured during activation by expression and activation by the 2 manipulations were in the lower portion of the face - the zygomaticus major (cheek raiser) and buccinators/risorius (lip retractor), the levator (nose region) and the orbicularis oris (surrounding the mouth). No muscles in the upper region of the face were measured because the manipulations were not expected to affect the upper region.

with motor simulation would be most detrimental to the recognition of happiness (Oberman, et al. 2007).

The motor activity associated with three manipulations intended to “disrupt” facial mimicry were compared. The three manipulations were: i) holding a pen horizontally between the lips and teeth without the teeth touching the pen or applying pressure (“hold”), as had been conducted by Niedenthal and colleagues; ii) biting down on a pen held sideways between the teeth without the lips touching the pen (“bite”); and, iii) chewing gum. Comparison of the muscles activity associated with each manipulation provided important information for planning future procedures for disrupting mimicry and understanding the impact of the manipulation on recognition. The most consistent motor activation was observed in the bite condition, equally affecting all four muscles measured. The chew manipulation activated muscles intermittently. The hold condition showed significantly less activity than the bite condition and no significant difference in activity between facial muscles (Oberman et al., 2007).

The three manipulations were applied during a categorization task. The bite manipulation caused the greatest interference with respect to recognition accuracy, the chew condition caused less interference, and the hold manipulation had no effect. The researchers reported that recognition impairment was related to overall muscle activity and not any particular muscle(s). Constant activation of the bite manipulation caused “interference” with muscles that would typically be activated during observation of a facial expression. The manipulation blocked spontaneous activity of muscles, making them unavailable to support processing the meaning of an expression, and primarily affected recognition of happiness and somewhat of disgust. According to Oberman et al., their findings support embodiment theories proposing that action recognition (facial expression) is supported by the same mental substrates that are engaged during production of the action.

Niedenthal and colleagues had observed impaired detection of change in expressions as a result of the “hold” manipulation. The results were explained in terms of congruence: facial activity generated by spontaneous mimicry supported an updated congruent state between the participant and the displayed expression. When free to mimic, change in the perceived expression triggered a change in the participants’ facial activity, generating updated feedback signals that were no longer congruent; this incongruence, in turn, facilitated rapid detection of change. Without spontaneous mimicry, detection of changes was impaired. The manipulation effectively delayed recognition of the change in expressions.

Like Niedenthal et al. (2001), Oberman and colleagues suggested that mimicry reflects the simulation of perceived expression in order to facilitate understanding. However, Oberman and colleagues did not observe an effect of the “hold” manipulation on a categorization task that required forced choice labeling of four very different expressions. They did observe impaired recognition accuracy during the bite manipulation, particularly with respect to happiness. Happiness was also the expression that generated the most facial activity, and the bite manipulation generated the most EMG activity relative to the hold manipulation. Although the intentions for the hold manipulation were to prevent motor activity (Niedenthal, et al., 2001) and thereby prevent smiling (Strack et al., 1988), the bite manipulation was intended to generate consistent and irrelevant “muscular noise”, making muscles unavailable for spontaneous mimicry (Oberman et al. 2007). Consequently, any relevant signal supporting simulation of the perceived expression was blocked.

These pivotal studies by Niedenthal et al. (2001) and Oberman et al. (2007) differed not just in terms of the task but also with respect to the quality of facial stimuli. One study required categorization of 500 ms black and white photos of prototypical expressions from the Ekman database, which had been modified somewhat for different levels of intensity. The other study required detection of change in 100 frame movies of morphed animations of facial expressions, which participants inspected frame by frame. Neither set of stimuli represented naturalistic expressions.

In a more recent study, naturalistic facial expressions were presented in videos as part of a task in which participants were required to decode true and false smiles (Rychlowska et al., 2014). In this study, participants wore a mouth-guard to inhibit facial motor activity. Participants who were free to mimic were significantly more accurate at discriminating genuine from false smiles compared to participants for whom mimicry was disrupted. Disruption of motor mimicry impaired participants' ability to recognize subtle differences in facial expressions and judgments about the meaning of smiles (Rychlowska et al., 2014).

Although the tasks and stimuli differed, each of the three studies illustrated an effect of disrupted facial mimicry that negatively impacted recognition and supported a functional role for mimicry in expression recognition. However, because of the inconsistent outcomes with the hold-pen manipulation, it remains uncertain what the effect of *inhibiting* mimicry really was. It appeared to have an effect on recognition of subtle differences in expressions or changes in expression, but no effect on a categorization task with expressions of slightly modified intensities. Differences between the hold and bite manipulations raised more questions about the effects of inhibition vs. interference with mimicry. Were effects on recognition due to a lack of essential motor feedback associated with spontaneous facial mimicry or due to active interference with common mental substrates that support activation and recognition of an expression? Thus, while mimicry clearly plays a role in expression recognition, the precise nature of that role remains unclear.

4.2 The current research

The main objective of this experiment was to test predictions regarding the functional role of spontaneous facial mimicry in social behavior in the natural world, where social interaction partners share a given set of circumstances or "situational context". To address this, an experimental approach was intended to measure recognition of an expression to a shared context. Specifically, as noted in the previous chapter, this approach included simultaneous presentations of a single static image of a context scene and a movie of a model's facial expression that either was or was not in response to the scene. Participants were required to make an explicit judgment ("true" or "false") regarding whether or not the model's facial expression represented an appropriate response to the given context scene.

With respect to the role of context, it was expected that recognition of the model's facial expression as a false response would be most accurate when the expression and context scene were of different valence (such as a negative facial expression of disgust in response to a positive scene of a grandpa playing with children). However, recognition of the model's expression as a true or false response was expected to be more difficult when the context scene and expression were of the same valence and same or similar in emotion.

Figure 4.1 shows examples of stimuli on these challenging same-valence trials. As reviewed earlier in this thesis, previous studies have shown that the inclusion of context influenced perception of emotion in static facial expressions, especially ambiguous expressions. Because of subtle differences in emotion between true and false responses on the same-valence trials, perception of the context scene was expected to influence processing of the facial expression: Participants' embodied response to the context scene was expected to influence their perception and subsequent judgment of the model's expression. With an underlying anticipation that the model's response would likely match their own response to a shared context scene, participants' perception and interpretation of the expression was vulnerable to the influence of the participant's embodied response to the scene. To that end it was expected that participants would show a tendency to judge the model's expression as being true, particularly when the context and facial expression evoked the same or similar emotions (i.e., the same valence). On these trials, it would be challenging to discriminate the expression as either the true or false response. As evident in Figure 4.1, in comparison, when the context scene and expression were of opposite valence emotions, recognition of the false response would be easy.

4.2.i Predictions for mimicry interacting with context

Hypotheses and predictions for mimicry and context were presented in Chapter 3. A possible interaction of mimicry with context is addressed here.

A facilitative role of mimicry that interacted with context should be evident on the difficult same-valence trials. It was predicted that spontaneous facial mimicry would reduce the influential effect of context on judgments of the expression as true, particularly on the more difficult same-valence trials. By strengthening the perceived expression of the model, mimicry would support discrimination of subtle differences between the perception of the model's response and the participant's embodied response to the context scene. Motor activity stimulated by perception of the facial response plus feedback from mimicry would boost simulation of the perceived expression, resulting in a more robust representation of the expression of the model. In light of previous evidence for early sensitivity to and detection of incongruence between perceived facial expressions and surrounding context (Righart & de Gelder, 2006), it was predicted that detection of incongruence between the represented expression of the model and the participant’s own embodied response to the perceived context scene would support recognition of the response as false. Such a facilitative role for mimicry was tested by disrupting mimicry. A stronger tendency to judge the model’s response as true when mimicry was disrupted would support the hypothesis for a facilitative effect of mimicry over the influence of context.

4.2.ii. Chosen manipulations for disrupting mimicry

While testing recognition of naturalistic expressions presented with context, mimicry was disrupted by means of the “bite-pen” manipulation used by Oberman et al. (2007) and the “hold-pen” manipulation used by Niedenthal et al. (2001). The impacts of these manipulations were compared.

Based on greater EMG activity generated by the *bite* relative to the *hold* condition (Oberman et al. 2007), the bite manipulation, by generating consistent irrelevant feedback noise, was expected to generate significant interference with the simulation process and therefore effectively disrupt the recognition process. Detection of incongruence would be difficult with constant feedback “noise”. Hence, it was predicted that while participants experienced the bite manipulation, spontaneous mimicry would be prevented, and if mimicry is beneficial to recognition, then the inference caused by the bite manipulation would impair recognition accuracy.

Consistent with the findings of Oberman et al., (2007), the *bite* manipulation was expected to have a stronger impact on recognition than the *hold* condition. The hold condition was intended to prevent mimicry and generate insufficient feedback (Niedenthal et al., 2001). With less motor activity and a weak, but still relevant, feedback signal, the participant's representation of the model's expression would be less robust. Nonetheless, the manipulation was found to alter recognition in other studies (Niedenthal et al., 2001); therefore, especially given the change in stimuli and procedure, it was worth testing this manipulation. No effect of the hold manipulation on situated recognition would indicate two possibilities: either the manipulation was not strong enough to prevent even a low amount of feedback that could benefit recognition or that facial feedback was not necessary for situated expression recognition. The latter would be confirmed if there was also no effect of the bite manipulation.

No effect of the bite manipulation would indicate that interference caused by irrelevant facial activity did not impact situated expression recognition. If disruption to mimicry did not affect recognition, then such an outcome would not support the simulated view of shared representations for perception and action as underlying the recognition process in the present paradigm. Alternatively, support for the simulated view would be found when disrupting activation caused disrupted recognition.

Methods

4.3.i Participants

Participants were 72 volunteers (M_{Age} = 20 yrs, ranging from 16 to 33 yrs, 92% female) All participants were undergraduate students at Dalhousie University, recruited through the Department of Psychology and Neuroscience volunteer subject pool and received 1% credit assigned to one psychology course. All participants declared normal or corrected-to-normal vision. According to the protocol approved by Dalhousie SSHREB, all participants signed an informed consent document. The *N* for this study matched the sample size of comparable studies testing for an effect of disrupted mimicry on recognition (Hess & Blairy, 2008; Maringer, et al., 2011; Niedenthal et al., 2001) and was expected to provide enough power to avoid a Type I error.

4.3.ii Materials

Context Scenes

A committee of 6 people (volunteers recruited through the Barresi lab at Dalhousie University, ages ranging from 21- 45 years), selected 32 scenes from the International Affective Pictures (IAPS) (Lang, Bradley & Cuthbert, 1999), a database of emotion eliciting scenes with standardized ratings for valence (positive/negative) and arousal (Lang, Bradley & Cuthbert, 1999). According to selection criteria, the scenes elicit eight emotions ranging from basic to more complex (anger, fear, disgust, sadness, amusement, excitement, contentment, and awe). Half the scenes had high ratings for eliciting one emotion in particular and half the scenes were rated for eliciting multiple emotions (e.g., sadness and anger) (Mikels et al., 2005). Sixteen scenes were sorted into similar emotion pairs and 16 scenes were sorted into contrasting emotion pairs. Figure 4.1 shows an example of context scene pairings. Table 2.1 in Appendix 1, provides a list of selected scenes, pairs and emotion categories.

Selecting the Models

Possible models were recruited through the volunteer subject pool for the Department of Psychology, according to a protocol approved by the Dalhousie Social Sciences Humanities Research Ethics Board (SSHREB), and received 1 bonus point towards a psychology course. Our goal was to find models that comprised cohorts with the anticipated sample of participants for later testing. Models were individually recorded while observing 40 context scenes, of which 32 were from our selected set. The extra scenes were incorporated as practice stimuli for later testing, or for possible replacements.

Seated approximately 60 cm from a 21.5-inch iMac G5 computer screen, facial responses were digitally recorded in real time at 29 frames per second; recording was done unobtrusively using the computer's built-in iSight camera while participants observed the scenes. Recording was programmed with an Apple script run on Apple Automator, so that recording of facial expressions occurred automatically during an 8 second (s) presentation of each scene. Following each presentation, recording stopped and the screen was black. Models were instructed to press the space bar on the keyboard to advance the program to the next scene once they felt their response had diminished. This self-monitored advance was intended to minimize contamination of expressions across responses.

To avoid possible order effects in responding, models saw one of two sets of the same IAPS scenes with different presentation orders. Models knew they were being recorded, were aware of our goal to collect facial responses to be used as stimuli in later experiments, and were instructed to "respond naturally" rather than posing or intentionally exaggerating (or inhibiting) their responses. The experimenter left the room after providing instructions to the model. Of the 24 models recorded, 4 model faces were selected (2 male, 2 female, mean age = 23 yrs) based on technical factors (no errors in recording), a visible range of responses, and equal representation of sex of model.

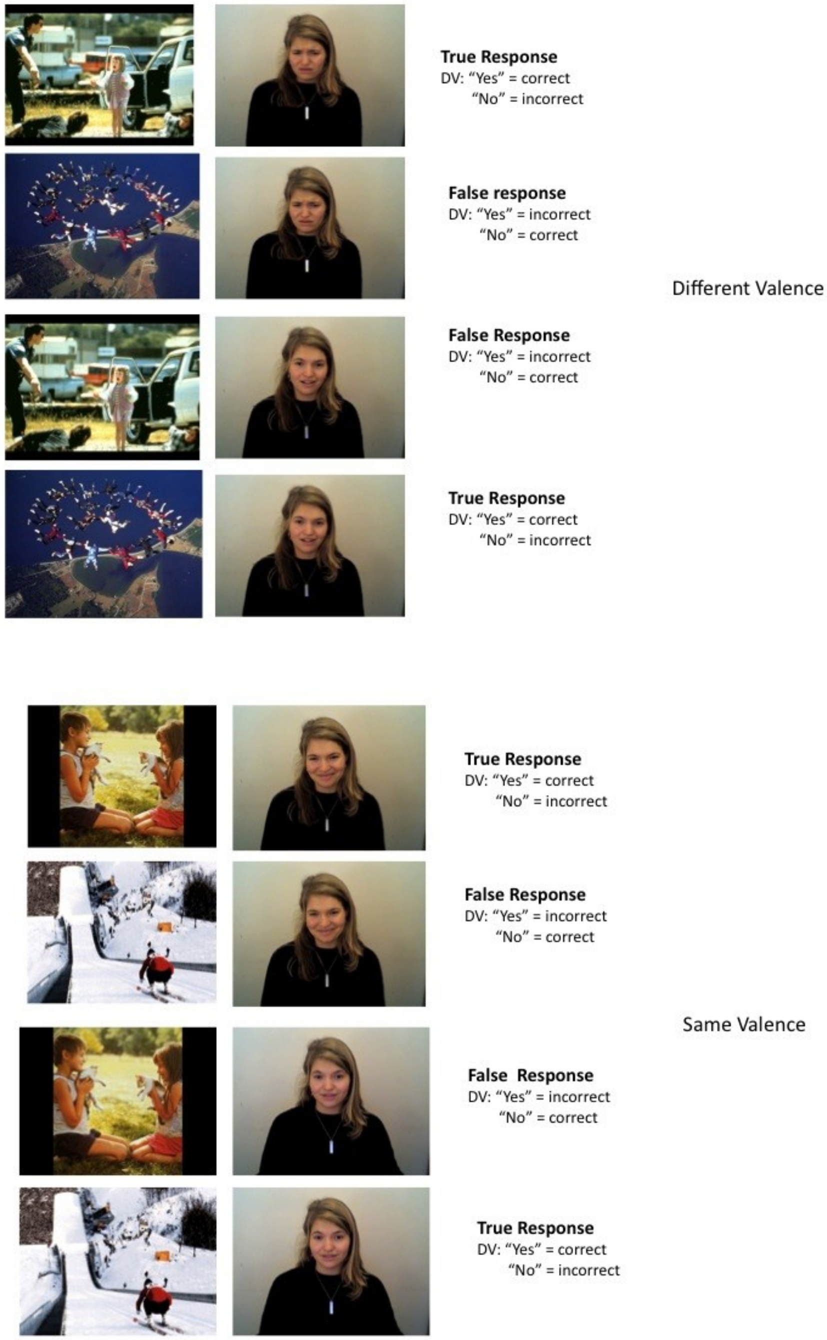

Figure 4.1 illustrates examples of one pair of different valence context scenes (upper) and one pair of same valence context scenes (lower) with the respective true and false response to each by one of four models. The "Yes" or "No" response is in answer to the question "Was this person looking at that image?" While the scenes above elicit fear, excitement and amusement, labeling of emotions was not required. The judgment is an attribution of the facial expression as the response to the context scene, or not.

Formatting the facial stimuli

From the 8 s of recorded facial response to each scene, a continuous 5 s clip was extracted using Final Cut Pro (v5.0.4.), saved as mpg files and later converted into mov format with a resolution of 480 x 640 pixels. There was no editing of the facial expression within the clip. Clip duration was five seconds, intended to allow for the capture of the natural rise and fall of an expressive response, or a possible change in expression. In addition to allowing for a change in expression, a 5 s exposure time was intended to enable processing of both the simultaneously presented context scene and the facial expression. Many recognition studies included unlimited exposure time controlled by the participant (e.g. Ambadar, Schooler & Cohn, 2005; Niedenthal et al, 2001) and although 5 s was longer than the exposure time of some recognition studies that measured facial activity (e.g. 500 ms, Oberman, et al., 2007) or rapid eye movements (e.g. 200 ms, Righart & de Gelder, 2008) the timing was consistent with previous behavioral studies (e.g. 4 s, Halberstadt, Winkielman, Niedenthal & Dalle, 2009; 3 s, Chan, Livingstone & Russo, 2009; 5 s, Moody, McIntosh, Mann & Weisser, 2007; 2040 ms, Sato & Yoshikawa, 2004) and shorter than other studies that used dynamic faces (e.g. 20 s, Davis, Senghas, Brandt & Ochsner, 2010; 15 s, Hess & Blairy, 2001).

There was a bank of 32 x 5 s responses identified by IAPS scene number for each of the four models. The naturalistic expressions were presented in a video format because, unlike static expressions, dynamic expressions presented as a movie allow for the detection of subtle emotions (Ambadar, Schooler & Cohn, 2005) and capture rapid shifts in focus and change in expression over time to different aspects of the situational context (McRorie & Sneddon, 2007). A range of expressions that were neither basic (the six described by Ekman) nor merely contrasting (e.g., happy vs. sad) was presented. Some expressions conveyed similar emotions (amusement and excitement), and some included a natural blend or shift in emotions (fear and anger). See Figure 4.1 for examples.

Stimulus presentation in the paradigm

Figure 4.2 shows the content of a single trial, which consisted of one static context scene (jpg format) and one 5 s movie of a model's response to a context scene (mov format, with programmed presentation time of 60 frames per second). The scene and movie were presented simultaneously and positioned side by side in the centre of an iMac G5 computer screen, with a screen ratio of 0,0, 1680,1050. Positioning of the face and context on the left and right alternated with trials.

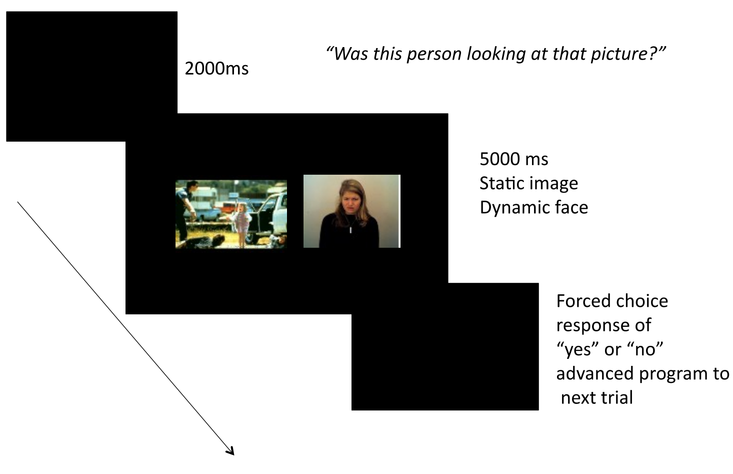

Figure 4.2 The procedure for a single trial.

Presentation of scenes and expressions was organized by the paired context scene; every scene was presented with a model response that was the actual response to that scene (the "true response"), and with a model response that was the response to the paired, foil scene (the "false response"). This meant that in 50% of trials the foil scene was presented with the false response that expressed a similar, same valence emotion; in the other 50% of trials the foil scene was presented with the false response that expressed an opposite valence emotion. (See Figure 4.1)

As Figure 4.3 illustrates, presenting true and false responses to 32 scenes by four models (F1,F2, M1, M2) totaled 256 trials. Stimuli were organized into two sets, with the presentation order of sets counterbalanced. As well as controlling for order, the sets controlled for which pairs of stimuli in the true and false conditions were experienced during disrupted mimicry. [7]

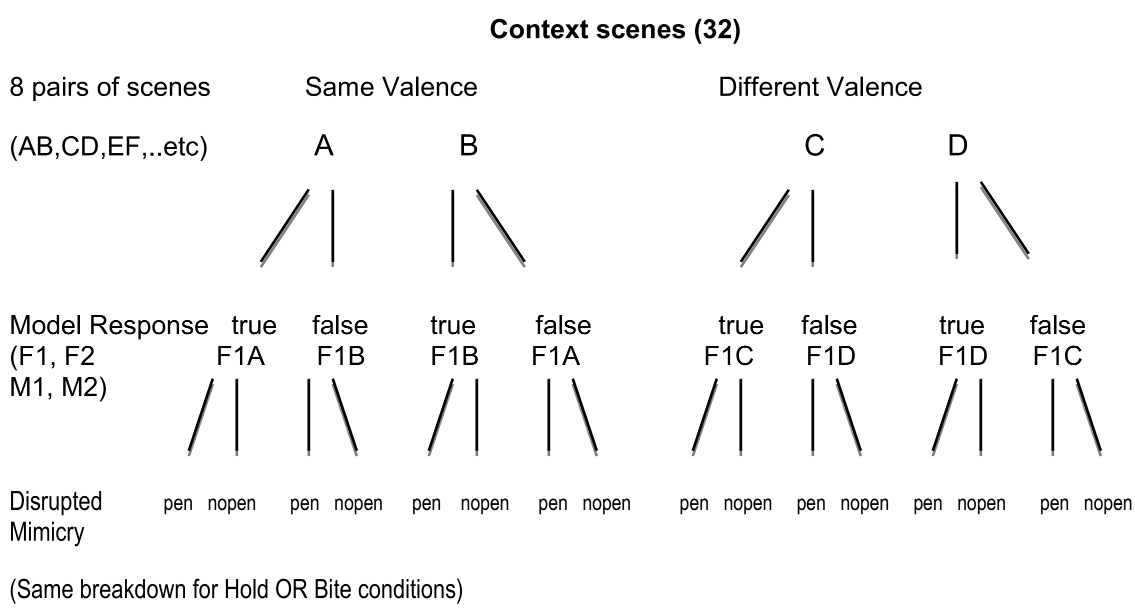

Figure 4.3. Organization of Stimuli: Thirty-two context scenes were sorted into 4 same-valenced pairs and 4 different-valenced pairs. The true and false response for each of 4 models (2 females, F1, F2 and 2 males M1, M2) was presented with the 32 scenes. Scenes are paired and for illustration purposes, coded here as A, B, C, D, etc. Model Response is coded here by model + scene letter, F1A is model F1's response to scene A. It is the true response to Scene A and the false response to scene B. During half the trials, mimicry was disrupted and during half participants were free to mimic. Half the participants experienced the hold manipulation and half experienced the bite manipulation.

[7] Additional information about testing the paradigm in an earlier study is outlined in the Appendix.

4.3.iii Experimental design

In a 2x2x2x2x2x4 mixed factor design, Model Response (true, false), Congruence of Context Valence (context scenes were organized into Same or Different Valence pairs), Disrupted Mimicry (50% of trials with the pen manipulation, 50 % without the pen), were within-subject factors. The condition of pen manipulation, Mouth Condition (bite, hold), control variables Pen1st (Pen 1st, Pen 2nd) and Presentation Order (1,2,3,4) were between-group factors. There were four blocks of 64 trials with a counterbalanced order of disrupted mimicry. Participants experienced one of two types of manipulations intended to disrupt mimicry, either the bite or hold manipulation. Over four alternating blocks of trials, half the participants experiencing the pen manipulation in the first and third blocks of trials and half experienced the pen manipulation in the second and fourth blocks. Participants were quasi-randomly assigned to either the bite or hold pen manipulation and to the Pen 1st or Pen 2nd groups. Presentation Orders of 1 – 4 were assigned sequentially.

4.3 iv Procedure

Programming of stimulus presentation was written by Chris Bebbington using Pygame modules for Python. For every trial, the dynamic face and static context scenes were presented simultaneously for 5 s. Immediately following offset of stimulus, the screen was black and participants responded to the question “Was the person looking at that scene?” by striking the “/” or “z” keys on the computer keyboard for “yes” (true) or “no” (false), respectively. Response keys were not counterbalanced. The ‘z’ key was consistently assigned for the “yes” response, with the intention of engaging the primarily non-dominant hand for the “yes” response so as to avoid a possible confound due to a dominant motor responses. The program did not proceed until a response was entered. A 2 s inter-trial pause followed the response. A break was programmed between two blocks of 128 randomized trials; participants were free to choose the length of this break, and continued to the next block by striking any computer key when ready.

Participation was conducted individually in a 10 x 18 metre (m) testing room in the Department of Psychology and Neuroscience at Dalhousie University. According to the approved protocol, the experimenter explained what a trial consisted of, as well as when and how to enter a response, and informed the participant of the emotion eliciting scenes that would be seen. Participants were told the purpose of the pen manipulation was to "learn if the pen would disrupt one's attention and thereby affect one's ability to interpret emotion in facial expression".

Following informed consent, participants were seated 50-60 cm from the iMac G5 computer screen and keyboard. Each participant was presented with 10 training trials consisting of 2 model faces (1 male, 1 female) that were not the same as the 4 testing model faces, and 5 context scenes that were not the same as the testing context scenes. From a box of new pens, participants selected a pen, which they could keep, and completed half the training trials with the pen in his or her mouth. They were instructed how to hold the pen in their mouth, pause the program and temporarily remove the pen if necessary due to discomfort.

4.4 Data analysis

Group means were initially submitted to a 2x2x2x2x2x4 mixed Analysis of Variance (ANOVA) with Manipulation ("bite/hold"), plus control variables Pen 1st or 2nd and Presentation Order as between-group factors. Disrupted Mimicry ("Pen/No-pen"), Congruence of Context Valence ("Congruent"/"Incongruent"), and Model's Response "True/False Response", were within-subject factors. Analyses of Variance were conducted using R version 2.15, with application of the "*ez* package", version 4.2.2

(Lawrence, 2009). A preliminary ANOVA was run on the full model that included control variables Presentation Order and Pen1st. Once it was determined there was no effect of Presentation Order or Pen 1st, these variables were dropped from the model.

Primary Variables of Interest

The dependent variable was Accuracy, defined as the proportion of correct responses. "Correct" responses (true or false) were correct identifications of the model's expressive response to the context scene. Each response was coded as 1 = correct or 0 = incorrect. The mean proportion of correct responses was calculated for every participant and data were organized into two groups defined by the pen manipulation to disrupt mimicry. Response times were not analyzed.

The variable "Model Response" corresponded to whether the model's expression was the true or false response to the context scene ("True", "False"). Accurate recognition of true and false model's responses was expected to vary between conditions when the context scene elicited an opposite valence emotion and when the context scene elicited a similar valence emotion. Recognition of the expression as a false response was expected to be easy when the context scene elicited an opposite valence emotion. In comparison, recognition of an expression as a true or false response was expected to be more challenging when the context scene elicited the same emotion or a similar emotion.

Context scenes were organized into pairs so that on a given trial the model's expression would be presented with the scene to which the display was the "true" response or with one other scene (the foil) to which the expression was the "false" response. Paired context scenes were either congruent or incongruent in valence. "Congruent-valence" pairs contained two scenes that expressed different yet similar emotions that were of the same valence (both positive or both negative) (e.g., one scene eliciting excitement and the other eliciting awe). "Incongruent-valence" pairs featured two scenes that expressed different emotions that were opposite in valence (positive and negative) (e.g., one scene elicited excitement and another elicited disgust). A variable named Congruence of Context Valence identified the context pairs as either "congruent–valence" or "incongruent–valence".

Participants were quasi-randomly assigned to one of two groups defined by the manipulations of either "holding" or "biting" on a pen in their mouth ("Manipulation" = hold = 1, n_{hold} = 35; bite = 2, n_{bite} = 36). The variable "Disrupted Mimicry" ("Pen", "No-Pen" identified trials with and without a manipulation that was intended to disrupt mimicry.

Secondary Control Variables

Application of the pen manipulation intended to disrupt mimicry was counterbalanced over alternating blocks. Forty-eight participants experienced disrupted mimicry with the pen manipulation throughout the 1st and 3rd blocks of trials ("Pen 1st") and 48 participants experienced the pen manipulation throughout the 2nd and 4th blocks ("Pen 2nd"). A variable named Pen 1st was added to the main model to determine if disruption to mimicry interacted with the order of experiencing the manipulation.

Presentation Order of stimuli was also counterbalanced. There were 4 subsets of stimuli, each with a single instance of a model-context pairing (32 images x True-False paired expressions), presented in one of four possible orders. Thus, another variable to control for order effects was labeled "Presentation Order" and added to the main model. Presentation Order was coded as; "1,2,3,4 = Order "1"; 2,3,4,1 = Order "2"; 3,4,1,2 = Order "3"; 4,1,2,3 = Order "4").

Alpha, effect sizes and outliers

The size of an effect was reported as general *eta* squared, which reliably measures the proportion of total variance in a dependent variable explained by the independent variable. Effect sizes calculated for general *eta* were expected to be smaller than the calculation for partial *eta* but are considered a more conservative measure for analysis of multiple factors (Bakeman, 2005). Alpha levels were set at a criterion of .05 for significance. Due to the multiple tests the risk of a Type I error was high. Therefore, Bonferroni corrections were applied. With 15 tests in the main model, the criterion for significance was adjusted to .003 (.05/15 = .0033). With 63 tests in the full model that included 2 control variables, the criterion for significance was adjusted to .0008. Dropping the control variables did not significantly change the F values in the main model of 15 tests. The F values between the larger ANOVA with control variables and the 15 test ANOVA for the main model were highly correlated (r = .999). Hence, except for the control variables, the results are reported from the main ANOVA of 15 tests.

Unless otherwise stated, paired t tests or one-way ANOVAs were conducted to test for simple effects. Errors bars represent Fischer's least significant difference, calculated using the mean N due to unequal groups.

A criterion for subject outliers for Accuracy was set at any subject mean that was greater than three standard deviations beyond the group mean. Subject means beyond that criterion rendered that subject an outlier and his/her data were dropped from analysis.

4.5 Results

The data from one participant (S166) were dropped from the hold condition because this participant's mean proportion of correct responses was below 3 SD from the sample mean (criteria for outliers are explained below). Thus, data from 71 participants were analyzed (n_{Hold} = 35, n_{Bite} = 36).

4.5.i Control variables and general findings

After Bonferroni corrections, the alpha level for the full ANOVA including control variables was .0008. There was no main effect of Order of Presentation, $F(1,55) = .0612$, $p = .97$, ${}_g\eta^2 = .000$ nor were there any two-way interactions. There was no main effect of Order of Pen ("Pen 1st"), $F(1,55) = .2200$, $p = .64$, ${}_g\eta^2 = .000$, and Order of Pen did not interact with Disrupted Mimicry, $F(1,55) = .4.674$, $p = .03$, ${}_g\eta^2 = .003$. Order of the Pen did not interact with Manipulation, $F(1,55) = .200$, $p = .66$, ${}_g\eta^2 = .000$. Order of experiencing the manipulations did not change the effect (or lack of effect) of disrupted mimicry on recognition accuracy.

There was a significant 3-way interaction of Order of Pen, Presentation Order, and Disrupted Mimicry, $F(3,55) = 20.744$, $p < .0008$, ${}_g\eta^2 = .04$. The interaction is explained by better performance on Blocks 1 and 3 than Blocks 2 and 4. This pattern of better accuracy over alternating blocks generated the appearance of a difference in accuracy due to Pen 1st and 2nd conditions and Disrupted Mimicry. However, across those two groups, the difference was explained by the presentation order of blocks 1 and 3. The only exception to this pattern was observed when the 3rd block of trials was presented first (In presentation order 3), when there was no difference in accuracy between blocks by participants who experienced the pen first pen.

Any further implication of the interaction required splitting the data by Time 1 and 2 and by Set. Because the way interaction of Order of Pen, Presentation Order and Pen was not significant with other variables this was not done and the interaction was not considered problematic.

The first set of results addresses the hypothesis of an influence of context on recognition. An influence of context on recognition of expressions was expected, particularly on difficult trials when emotions in the expression and context scene were similar (of the same valence). It was predicted that the influence of the participant's embodied response to context would manifest as a tendency to judge the response as congruent with one's own and therefore as a True rather than False response. In other words, a response bias to say 'True' was predicted. Thus, higher levels of recognition accuracy were expected for True relative to False responses.

The second part of the analysis addressed the effect of disrupted mimicry on recognition of situated naturalistic expressions. According to the hypothesis that mimicry facilitates recognition of facial expressions, disruption of mimicry was predicted to interfere with recognition accuracy, particularly when the model's response and context scene were of similar emotions. It was expected that participants' facial activity generated by mimicry would be detected as congruent or not with the participant's embodied response to the context scene. Incongruence would be indicated by a change in representation that differed from the embodied response. Furthermore, one hypothesis was of mimicry as a mechanism that would reduce the effect of context by accentuating subtle differences between the mimicked model's response (via feedback) and the participant's embodied response to the context scene. It was expected that when participants were free to mimic they would be equally accurate recognizing True and False responses. Conversely, disrupted mimicry was expected to result in a more pronounced influence of context on the participants' judgment. Participants' discrimination of any difference between their embodied response to context and the mimicked response to the expression would be compromised with disrupted mimicry, and the participant's embodied response to context would dominate their processing of the model's response. Therefore, disrupted mimicry was expected to result in a greater tendency to judge the response as True (as a match to their own response to the context) and greater recognition accuracy to True than False responses. There were two conditions of manipulation intended to disrupt mimicry: the bite and hold conditions.

4.5.ii Tests for an influence of context on judgments of model response

The overall mean proportion of correct responses for the total sample was .70(*SD* = .05). This mean was statistically greater than chance, $t_{(70)}$ = 33.71, *p* < .0001, *SE* = .006, *Mean Difference* = .20, CI = 0.1882 – 0.2118. There was a robust, main effect of Congruence of Context Valence on accuracy, *F*(1,69) = 450.922, *p* <.0001, ${}_{g}\eta^2$ =.26. Collapsed across the between-group factor of Manipulation, all participants were significantly more accurate on trials in which the context scene was of a pair (the actual and foil) that were incongruent in valence, ($M_{\text{Incongruent}}$ = .78, *SD* = .06), compared to trials in which the context scene was of a pair that were the congruent in valence ($M_{\text{Congruent}}$ = .62, *SD* = .05). The size of this effect explained 26% of the total variance in recognition accuracy. Hence, the anticipated difference in difficulty between trials involving congruent- and incongruent-valence context pairs was found.

The importance of the effect of Congruence of Context Valence was best understood by the interaction with Model's Response. Figure 4.4 shows that, overall, participants were no better at recognizing the Model's Response as True than False, *F*(1,69) = .519, *p* = .47, ${}_{g}\eta^2$ =.002; however, a strong interaction between Model's Response and Congruence of Context Valence was observed across all participants, *F*(1,69) = 223.744, *p* < .0001, ${}_{g}\eta^2$ =.145. With the size of the interaction accounting for more than 14% of total variance in accuracy, congruence of the context scene was clearly an important factor influencing judgment of the model's response as true or false.

As expected, recognition of the true response was equally challenging regardless of whether context scenes were of congruent or incongruent-valence pairs. Lower levels of accuracy on congruent valence trials were explained by this difficulty of distinguishing same from similar emotions regardless of their pairing. In comparison, and as expected, recognition of the false response was difficult only when the foil scene elicited congruent valence emotions, but easily detected on trials in which the context scene and model response expressed incongruent valence emotions. When discrimination of the true and false response was difficult, participants showed a greater tendency to judge the response as true. The interaction explains why the main effect of Model's Response was not significant. The difference in accurate recognition of the true and false response was apparent only in relation to the valence of the context scene.

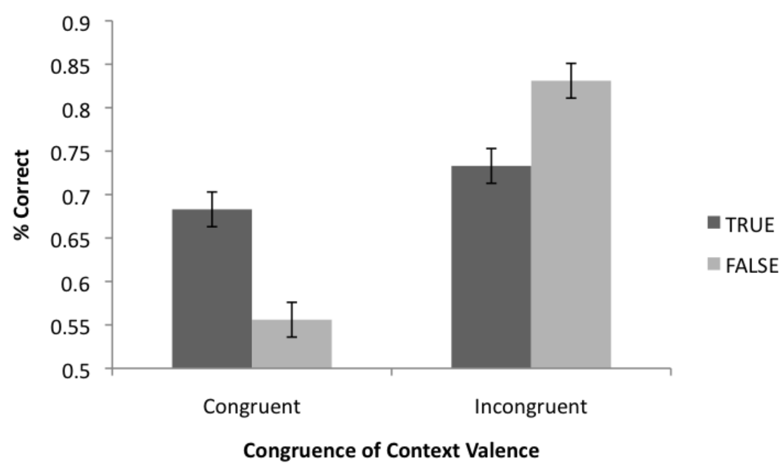

Figure 4.4 A two-way interaction of Model's Response (True/False) and Congruence of Context Valence (congruent/incongruent) was significant ($p < .0001$, ${}_{g}\eta^2 = .15$). Accuracy was significantly lower to false responses when context scenes were congruent in valence relative to when context scenes were incongruent in valence. $N = 71$. Error bars depict *FLSD*.

The magnitude of the effect of context on recognition of the true and false response supported the hypothesis of an influence of context on expression recognition. Discrimination of the true and false response was difficult when context scenes were congruent in valence. As expected, when faced with uncertainty, processing the expression was influenced by the participant's embodied response to the context scene, as well as an expectation that the other person's response to the shared context would match their own. Consequently, on these difficult trials, participants were more likely to judge the expression as "true". The exception to this effect was, of course, easy recognition of the false response when the foil context scene expressed an incongruent (opposite) valence emotion. Thus, the hypothesis for an influential effect of context was supported, particularly when the model's response and context scene expressed the same or similar emotions. This effect was also confirmed with criterion analysis. A liberal bias to judge the model's response as true was confirmed on congruent trials, p = .0004.

4.5.iii Tests of effects of manipulations intended to prevent spontaneous facial mimicry on accuracy

Another main objective of the analysis was to determine if mimicry supported recognition of situated naturalistic expressions. There were no main effects of Disrupted Mimicry on accuracy, $F(1,69) = 1.765$, p =.19, ${}_{g}\eta^2 = .008$, as participants were no less accurate when mimicry was disrupted ($M_{\text{Pen}} = .71$, $SD = .05$) than when they were free to mimic ($M_{\text{NoPen}} = .70$, $SD = .05$).

This pattern of results for Disrupted Mimicry was the same for the hold and bite groups. Based on reports by Oberman et al., (2007) there was an *a priori* expectation of greater impairment to recognition caused by the bite than the hold manipulation. However, there was no main effect of the Manipulation that affected Accuracy, M_{Hold} =

.71, SD = .05, M_{Bite} = .70, SD = .05, $F(1,69)$ = 1.569, p =.21, ${}_{g}\eta^2$ =.014. Disrupted Mimicry itself did not interact with Manipulation, $F(1,69)$ = .206, p =.65, ${}_{g}\eta^2$ =.001. Regardless of the type of manipulation, disruption to mimicry did not impair recognition. In this first experiment, the hypothesis of mimicry underlying expression recognition was not supported.

Furthermore, Disrupted Mimicry did not interact Congruence of Context and Model Response to alter the influence of context. The three way interaction from the main model was not significant, $F(1,69)$ = 1.235, p =.27, ${}_{g}\eta^2$ =.000. Nor was there a significant interaction of Manipulation with Congruence and Model Response, $F(1,69)$ = .040, p =.84, ${}_{g}\eta^2$ =.000. Figure 4.5 summarizes how an influence of context did not change when participants experienced the bite or hold manipulations that were intended to disrupt mimicry. Regardless of manipulation, levels of accuracy were higher for judgments of the response as true compared to false when the context scene was congruent in valence, but not when context scenes were incongruent. Figure 4.5 also shows that, in contrast to the *a priori* expectation of a difference in effect of the bite and hold manipulations, there was no difference in accuracy caused by either manipulation. Because of the *a priori* expectation for a stronger disruptive effect of the bite compared to the hold manipulation on judgments, select analyses were run on the data for each group. Disrupted mimicry accounted for .00% variance in accuracy for both the bite and hold groups. The effect of context indexed by the interaction of Congruence of Context valence and Model Response was stronger on judgments by the bite group compared to the hold group. Context accounted for 21% variance in accuracy when mimicry was disrupted by the bite manipulation compared to 18% variance in accuracy when mimicry was disrupted by the hold manipulation.

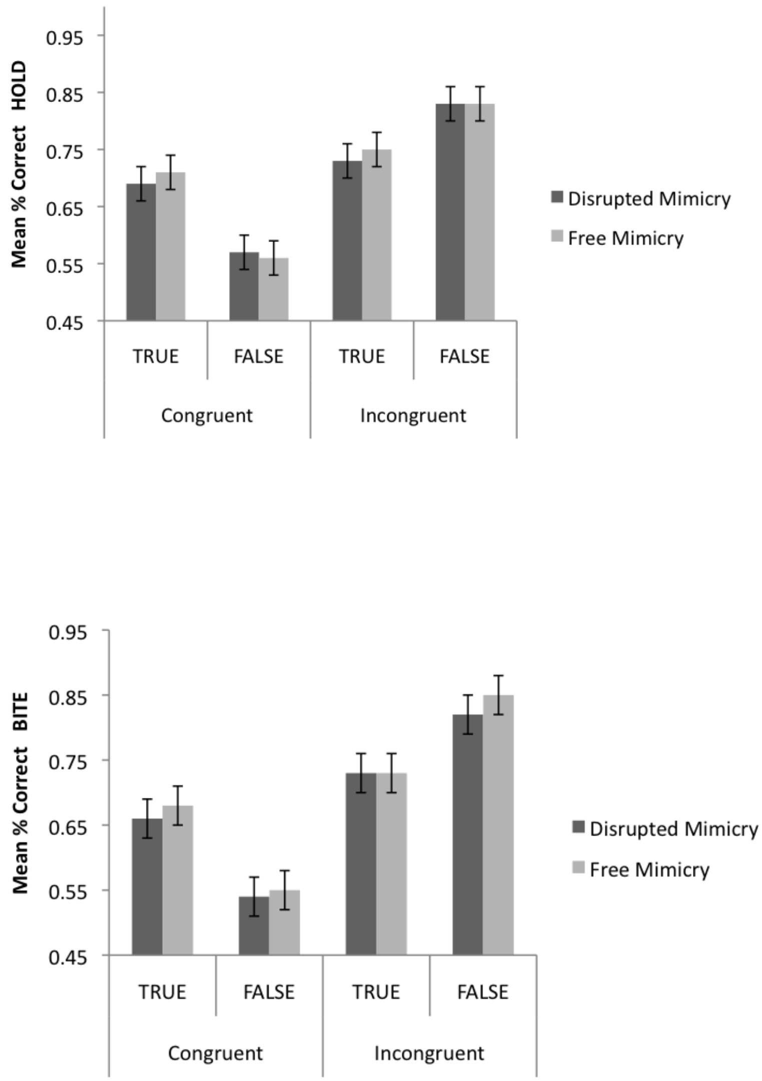

Figure 4.5 When participant experienced the Hold (upper) or Bite (upper) manipulation disrupted mimicry did not interact with Congruence of Context Valence and Model Response to change the influence of context. n_{Hold} = 35, n_{Bite} = 36 Error bars are FLSD.

4.6 Discussion

Predictions for an influential effect of context on participants' judgment of the model response were strongly supported. When the context scene and model's response expressed the same or similar emotions (which was in all cases except the false response to a context scene of an opposite valence emotion), participants were more accurate when the model's response was true than when it was false. High levels of recognition accuracy to true responses were explained by participants' tendency to judge the model's response as true, resulting in poor recognition accuracy of the false response on same-valence trials. Only when the context scene was obviously expressing an opposite valence emotion were levels of recognition accuracy high, as false responses were easily detected. Otherwise, on all other trials, participants showed a tendency to judge responses as true, which reflects the effects of context on judgment. When discrimination of the model's response was uncertain the participant's embodied response to the context scene influenced their perception and processing of the model's response. This finding was consistent with reports of the influence of context on perceptual processing of facial expressions (Aviezer et al., 2008; Halberstadt et al., 2009; Hassin, Aviezer, & Bentin, 2013; Masuda et al., 2008; Righart & de Gelder, 2006; 2008).

For instance, according to previous reports in the literature, observers' responses to context influenced their perception of emotion in faces that were embedded onto cutout figures (Aviezer et al., 2008; Hassin, Aviezer, & Bentin, 2013). Eye-scanning patterns were altered in response to a change of face-figure combinations. This change in processing of the facial information was interpreted as evidence of a modulating effect of context on perception occurring at an early, bottom-up level of processing (Hassin, Aviezer, & Bentin, 2013). In the same way, in the current work, the participant's embodied processing of the context scene likely influenced their early attention and perception – and subsequent interpretation – of the model response.

Other studies reported a top-down effect of context on face perception. Emotion concept labels presented with facial expressions influenced encoding of facial expressions. Participants who initially perceived faces with positive emotion labels later rated the same faces as more positive than did participants who had seen the faces without emotion labels (Halberstadt et al., 2009). Processing of emotion concept words influenced initial encoding of the faces and later perception of emotion in the same facial expressions.

These previous reports support the interpretation of findings in the current study; that processing of the context image influenced participants' processing of the expression and consequent judgment of the model's response. Unlike other studies in which the inclusion of context was implicit, in our paradigm participants were required to explicitly consider the context. In the current study, explicit processing of the context scene, plus an implicit priming effect of the observer's response to the scene on their processing of the expression, helped to explain the tendency for participants to judge responses as "true".

Consistent with theories of simulation, including those pertaining to embodied cognition, in the current work it was proposed that active consideration of the context scene contributed to a robust embodied processing of the information. When information in the perceived expression was possibly ambiguous (difficult to discriminate as the true or false response because both contain similar features that express a similar or same emotion), the robust embodied response to context likely primed perception of the ambiguous expression. Such priming resulted in the participant's interpretation of the ambiguous model’s response as matching his or her own and therefore representing the true response to the given context.

Mimicry was another primary variable of interest. Consistent with the hypothesis that mimicry supports recognition of expressions, recognition was expected to be more accurate when participants were free to mimic than when mimicry was disrupted. Mimicry was expected to facilitate recognition by contributing information about the perceived emotion of the model, via facial feedback, that would be congruent or incongruent with the participant’s embodied response to the context scene. Similarly, it was also suggested that mimicry of the model’s expression generated increased levels of motor activity that contributed to the participant’s simulation of the perceived expression. Detection of congruence between the mimicked model’s response and participant’s embodied response to context was expected to facilitate a judgment of the model’s response as true. Conversely, when feedback generated by spontaneous mimicry of the model’s expression was not congruent with the participant’s response, incongruence would be detected. That detection of change was expected to facilitate recognition of the model’s response as false. Thus, mimicry would moderate the influence of context on participants’ judgment and conversely a disruption to mimicry would result in a greater influence of context.

These expectations were partially based on prior reports of disrupted mimicry resulting in impaired recognition of expression (Livingstone et al., 2016; Maringer et al., 2011; Niedenthal et al., 2001; Oberman, et al., 2007; Rychlowska, et al., 2014). Until the study by Rychlowska et al. (2014), no other known studies had measured how disrupted mimicry would interact with context to affect emotion recognition on naturalistic expressions.

Predictions for impaired recognition as a direct result of disrupted mimicry were not supported. Participants were not more accurate in their judgments when they were free to mimic than when spontaneous mimicry was disrupted. Based on the findings reported by Oberman et al. (2007), there was an *a priori* expectation of a stronger effect of the bite relative to the hold condition in disrupting mimicry. This difference was not found. Responses by participants in the hold and bit conditions showed no significant effects of either manipulation.

According to Niedenthal et al. (2001), the manipulation in the hold condition inhibited overt facial mimicry and, while the facial activity generated by the manipulation itself was minimal, low levels of facial mimicry in areas other than the mouth were possible (as reported by Rychlowska et al., 2014). The manipulation wasn't strong enough to prevent facial feedback contributing to the simulation of the perceived expression. In the current work, the hold condition was considered no different from a neutral condition when participants were free to mimic.

According to EMG measures reported in previous studies (Oberman et al., 2007), the hold condition generated significantly less facial activity than the bite condition. Imaging studies have also shown activity in both the primary and premotor areas during imitation and observation of facial expressions; and during passive observation only, activity in premotor area was more robust than activity in the primary motor areas (Carr, et al. 2003). Thus we might infer that the inhibitory effect of the hold manipulation during observation of expressions interfered with primary motor activity but that the strength of the signal caused by the manipulation was likely not strong enough to interfere with premotor activity. The signal from the bite manipulation, however, was likely strong enough to interfere with both primary motor and premotor cortical activity, contributing to a change in processing for recognition of the expression.

According to Oberman et al. (2007), the bite manipulation prevented spontaneous mimicry, at least for expressions that relied on activity in the lower portion of the face. With activation of the manipulation that was intended to disrupt mimicry, participants' processing of the model's expression was expected to be compromised: constant activation of the zygomaticus major and buccinator muscles that typically support smiling made them unavailable for spontaneous activation that would, under normal circumstance, take place during processing of the expression. The signal noise generated by that consistent activity was irrelevant to the perceived expression and interfered with participants' representation of the perceived model's expression. In the current work, the interference effect was not found on judgments about naturalistic, dynamic expressions that were presented with context.

It was predicted that spontaneous mimicry would moderate (and reduce) the effect of context, such that there would be no difference in accuracy between true and false responses when participants were free to mimic. In keeping with the results of Maringer et al. (2011) and Rychlowska et al., (2014) it was expected that participants would show greater ability to detect fine grain differences in the expressions when they were free to mimic and therefore, in this experiment, recognize when an expression was the true or false response. In contrast, a greater tendency to judge subtle expressions on congruent valence trials as true was expected when mimicry was disrupted. That was not observed. Figure 4.5 illustrated no difference in accuracy due to disrupted mimicry in either conditions of manipulation.

In an effort to understand the lack of effect of the bite or hold manipulations, which differed from previously published effects, differences in experimental design between the current work and procedures of key comparative studies were considered. In striving for ecological validity, the current paradigm differed in several ways related to task, context, and quality of facial stimulus. The labeling paradigm conducted by Oberman et al. (2007) did not include context, and tested for an effect of disrupted mimicry on black and white, static facial prototypes that were presented for 500 ms. Niedenthal et al. (2001) incorporated context in the form of emotion words, and tested recognition of change from positive to negative and negative to positive expressions in computer-generated movies with no time limit. Participants controlled the speed of presentation and examined each expression frame by frame. In so doing, their examination was essentially on a series of static images. In comparison, the situated recognition task in the current work required explicit consideration of a static context scene and judgment of dynamic naturalistic facial expression as the true or false response to the given context scene.

Of particular interest to this research are more recent findings reported by Rychlowska et al (2014) who measured the effect of blocked mimicry (by having participants wear mouth guards) on accurate recognition of genuine smiles in movies of naturalistic smiles. In their study, interference with the ability to freely mimic resulted in reduced ability to discriminate fine details in the smile. The recent work by Rychlowska et al. (2014) most closely resembles the current work because of their use dynamic, naturalistic facial stimuli and the need for participants to discriminate subtle differences in similar expressions.

Dynamic expressions are rich stimuli that provide information not available in the static presentation of a facial expression. For instance, temporal spatial information and perceived motion is readily available in the dynamic expression (Ambadar et al., 2005). These may be important qualities missing in the static expression and that are generated by facial mimicry. Furthermore, there is accumulating evidence for differences in neural systems that support the processing of dynamic and static expressions. Neuroimaging studies have identified different levels and regions of activity for processing static and dynamic emotion faces (Arsalidou, Morris & Taylor, 2011; Kilts, Eage, Gideon, & Hoffman, 2003; Trautmann, Fehr, & Herrmann, 2009). The evidence strongly indicates widespread activation of multiple systems for processing dynamic faces, and less widespread areas of activity but concentrated and sometimes higher levels of activity in the motor and premotor systems for processing static faces. In light of these neuroimaging findings, mimicry may be beneficial but not necessary to processing dynamic expressions. On the other hand, with fewer systems actively supporting processing of static expressions and one of those systems being motor (and premotor), then facial mimicry may be particularly beneficial for processing static expressions. Thus, the *disruptive* effect of the bite manipulation on a labeling recognition task, as reported by Oberman and colleagues (2007), may have been specific to recognition of static rather than dynamic expressions. Consequently, the interpretation of the evidence garnered by previous studies using static faces may be limited to perception of photographs and may not adequately represent interpersonal interaction in the natural world. However, more recent work has indeed shown reduced levels of accuracy to dynamic smiles when mimicry was blocked (Rychlowska et al. 2014). Clearly, there is more to be learned about the role of spontaneous mimicry in *situated* emotion recognition. The next study in this thesis compared the effect of disrupted mimicry on static as well as dynamic naturalistic expressions.

Thus, the goal of the next experiment was to achieve a better understanding of under which conditions disrupted mimicry interferes with situated recognition, if indeed it does. The bite manipulation was shown by Oberman et al. (2007) to have a stronger effect than the hold manipulation and to interfere with spontaneous activation of the zygomaticus major and buccinator muscles, which typically support perception and production of smiling. Consistent with Oberman et al.'s interference hypothesis, the bite manipulation should cause a differential effect on judgments of facial expression in response to positive and negative stimuli.

Controlled application of only congruent valence context scenes in a subsequent experiment would help to clarify how the manipulation intended to disrupt mimicry affected judgments of positive and negative expressions. Furthermore, context scenes that were congruent in valence would test for a replication of the robust effect of context observed in this first study. A replication would support the embodied perspective from stimulation theory that suggests the observer's embodied response to a shared context influences perception and judgment of another person's facial expression.

CHAPTER 5: EXPERIMENT 2

The rationale for testing situated expression recognition of *dynamic* facial expressions was largely based on the objective to increase ecological validity. Dynamic expressions reflect our everyday social interactions: Natural expressions are inherently dynamic, revealing constant changes in underlying emotions and motivations that an individual experiences in response to their social situation. Previous use of static expressions in recognition studies had been criticized as having low ecological validity. Natural expressions are dynamic not static, they are spontaneous and not intentionally posed to meet a set criterion for a given emotion, nor are they artificially generated or manipulated. Computer generated expressions successfully convey change in an expression, but the motion is artificially produced by manipulating a series of static images. Thus, dynamic naturalistic facial expressions that were representative of daily experience were presented in the task for this research.

Both static and dynamic expressions contain structural information from which the basic configuration of an expression is perceived, yet dynamic expressions are purported to be more complex and salient compared to static facial expressions. Increased complexity and saliency is likely due to the rapidly changing sequence of information revealed in the motion of the expression (Ambadar, Schooler, and Cohn, 2005). Moreover, it has also been suggested that constant motion cues in a dynamic expression capture attention and increase arousal in the observer (Trautmann, Fehr and Herrmann, 2009). There is a growing body of research identifying different outcomes for emotion recognition in static and dynamic expressions that led to a hypothesis in the literature of separable but overlapping systems for processing static and dynamic facial expressions (Kilts, et al., 2003; La Bar, 2003; Pitcher, et al., 2012;Trautmann, et al, 2009; 2013). The following review of behavioral and neuroimaging research supports an argument for a difference in systems that actively underlie processing dynamic and static expressions.

Impaired recognition had been previously reported for the two pen manipulations applied in Experiment 1, both of which were intended to disrupt mimicry (Oberman et al 2007; Niedenthal et al., 2001). However, in contrast to those reports, the results for Experiment 1 showed that neither manipulation caused significant impairment in recognition. The dynamic quality of facial stimuli presented in the current research may explain the unexpected lack of effect. Situated dynamic presentations of spontaneous, naturalistic facial expressions were presented in Experiment 1 as opposed to static, prototypical, black and white images of facial expressions (Oberman et al., 2007) or movies of computer generated facial expressions morphed from black and white images of prototypes (Niedenthal et al., 2001). Reasons for presenting dynamic expression were based primarily on striving to increase ecological validity.

Thus, it was decided to continue testing the hypothesis of mimicry as underlying expression recognition (Atkinson & Adolphs, 2005, Oberman et al., 2007; Niedenthal et al 2001). Impaired recognition resulting from the disruption would support the hypothesis. Impaired recognition of static (as per reports in the literature) but not dynamic expressions (as per results in Experiment 1) would suggest that mimicry underlies recognition in static but not dynamic expressions.

Assertions presented in the literature for “separable systems” for recognition of static and dynamic facial expressions (Kilts et al., 2003) were based on findings from neuroimaging studies comparing patterns of neural activity in response to static and dynamic expressions (Arsalidou, Morris & Taylor, 2011; Kilts et al, 2003; La Bar et al., 2003, Pitcher et a. 2011; Trautmann et al., 2009; 2011). Different outcomes for a disruption to mimicry on recognition of static and dynamic expressions in the current behavioral study would support assertions for a difference in underlying processes for recognition of dynamic and static expressions. Moreover, different outcomes in a *situated* expression recognition task that involves a judgment about an expression as appropriate to a given context would also suggest a difference in the how the processing systems interact with context.

Three research questions were addressed in the current study. One question asks if spontaneous facial mimicry underlies recognition of static *and* dynamic expressions. A second question asks if the inclusion of context in the situated expression recognition task would differentially influence judgments about dynamic and static expressions. A third question asks if there is a dynamic advantage to expression recognition. As reviewed below, reports of better accuracy to dynamic relative to static expressions were frequent in the literature.

5.1 Is there a dynamic advantage to expression recognition?

Different outcomes for testing recognition of expressions that were static or dynamic were the first indicators of different processes for the quality of stimuli. There was accumulating evidence for a dynamic advantage in recognition of emotion in dynamic relative to static facial expressions (Ambadar et al., 2005; Bassili, 1978, 1979; Back, Jordan and Thomas, 2009; Bould & Morris, 2008; Hoffman et al., 2013; Kilts et al. 2003; Krumhuber et al. 2013; LaBar et al., 2003; Trautmann et al, 2009, 2013; Weyers, Muhlberger, Hefle & Pauli, 2007). There have been, however, inconsistencies in results and the conditions for different outcomes have been debated. Quality of facial stimuli has been proposed as one explanatory factor for the inconsistencies. For instance, Fiorentini and Viviani (2011) contrasted recognition accuracy to two qualities of facial stimuli; 'realistic' blends of two dynamic expressions and static versions of the apex of the blend. They did not observe a difference in accuracy to dynamic and static expressions and suggested that a so-called *dynamic advantage* to emotion recognition in facial expressions was more likely to occur when the quality of faces was degraded and not realistic. Consistent with this suggestion, superior recognition of dynamic expressions was indeed reported when all stimuli were degraded [e.g. due to poor illumination (Bassili, 1978, 1979), wire construction (Krumhuber et al., 2013), or point light faces (Atkinson, 2007)]. However, a dynamic advantage for recognition accuracy was also reported for natural, fully illuminated expressions (Bassili, 1978, 1979) and natural expressions revealing complex mental states (e.g. anxious, suspicious, and disapproving, reported by Back, Jorden and Thomas, 2009). Thus, the argument for a dynamic advantage explained by degraded quality of static stimuli was not fully supported. Nor did a range of quality of dynamic images compromise a dynamic advantage.

Computer morphed animations built from static images of facial expressions generate an artificial dynamic stimulus. Morphed animations of facial stimuli can be manipulated to control for factors of interest such as intensity of an expression, velocity (the speed at which the expression unfolds), change from one expression to another, and activity in selected facial areas of interest. Although not natural, these computer generated dynamic expressions are not typically considered “degraded” (unless intentionally generated to be so). For example, Fioretini and Viviani (2011) presented morphed animations as ‘realistic’ versions of expressions. Common application of these stimuli in recognition studies has not hindered the dynamic advantage to accuracy (Ambadar et al. 2005), nor has the application of avatars – an entirely computer generated facial stimulus (Weyers, et al, 2007). Ambadar and colleagues (2005) reported a dynamic advantage to *subtle* expressions that were generated by morphing static images. Their dynamic expressions emerged from neutral and progressed until the first display of the expression was visible. Recognition was better in their dynamic condition relative to their single static or multi-static conditions.

Although computer morphed animations of dynamic expressions are not considered degraded, the degree to which they appear natural has varied across studies. For instance, faces presented by Ambadar et al. (2005) were generated by morphing techniques and appeared realistic in that they were not avatars, but the unfolding of their dynamic expressions was described as 'synthetic', and therefore not natural. Multi-frame "dynamic sequences" unfolded over 4-6 images that were held for different times (i.e. the first neutral image was held for 500 ms, then 2 to 4 images of intermediate expressions held for 33 ms and a final image held until participants responded). Thus, although their dynamic expressions were intended to be more representative of those encountered in daily activity, relative to static images (Ambadar et al., 2005), the unfolding was not entirely natural. Additionally, in their stimuli, an oval surrounding the face, masking features such as hair, neck and shoulders, hindered the naturalness of their presentations**.** Therefore, the evidence existed for a dynamic advantage, but with artificially generated dynamic. It was not known if the advantage would generalize to natural stimuli.

Possible factors to explain a dynamic advantage

Ambadar et al. (2005) tested the hypothesis of the benefit to recognition accuracy afforded by dynamic faces was an *increased amount of information* relative to a static facial expression. They questioned if the dynamic image of a facial expression was equivalent to multiple static images. In other words, was the increased information in a dynamic expression equivalent to "additional static" information. Different versions of subtle facial expressions - dynamic, static, multi-static – were presented in a recognition task. The purpose of the multi-static version was to weaken the effect of motion so that an effect of ‘additional static information’ could be measured. Thus, the multi-static version included visual noise masks inserted between single frames used in the dynamic sequence. The researchers reported better recognition accuracy with the dynamic version than the static and multi-static versions and concluded that a dynamic advantage could not be explained by additional static information.

Instead, according to Ambadar et al. (2005), information perceived in their dynamic expression supported the *perception of change* that was not available from perception of a static image, or from multiple static images of facial expressions. The notion of “change” in a dynamic expression suggests a transformation and involves capturing a range of expressivity. Unlike a static image of an expression (which, as repeatedly acknowledged, are typically built from the apex of the expression) that has limited, non-changing information, a dynamic expression typically reveals a transformation of an expression as it emerges from neutral to its peak. It was the “perceived change” and range of expressivity available in the dynamic expression that contributed to better recognition accuracy (Ambadar et al., 2005).

By replicating the methodology used by Ambadar et al. (2005) researchers contrasted recognition of *subtle* and *intense* static and dynamic expressions (Bould & Morris, 2008). A dynamic advantage was replicated on subtle expressions, but not on expressions of greater intensity. Bould and Morris (2008) concluded that when the static image contained the expression at its apex there was no additional information required for recognition that might be available in the dynamic expression. Additionally, the dynamic sequence presented by Bould and Morris was modeled after the stimuli of Ambadar et al. (2005), and although their images were not generated as morphed static images, the dynamic sequences were not ‘natural’. Their dynamic stimuli contained 4 – 5 non-sequential still frames selected from a recorded facial expression. To show the emergence of a subtle expression the 4 frames were run together as a moving sequence. The first neutral frame was held for 500ms, 2 intermediate frames for 200 ms and the final frame for 500 ms. Their “intense” version of expressions contained one extra intermediate frame and the final image was held for longer (1000 ms). The researchers acknowledged that the natural unfolding of an expression was lost in their dynamic condition (Bould & Morris, 2005). Although their argument for a dynamic advantage in subtle but not intense expressions was compelling, their conclusions were based on artificial dynamic expressions, and were therefore limited.

Contrary to Bould and Morris’s conclusion regarding the sufficiency of an intensely expressive static expression, difficulty in discriminating static expressions has been documented. For instance, static images of fear and surprise expressions are difficult to discriminate because eyes wide open could signal either fear or surprise (Hoffman et al, 2013; Sato & Yoshikawa, 2004). Yet, by controlling selective areas of facial activity in morphed expression, emotion-specific effects of motion were found; e.g. recognition of fear was better with motion in the upper face area relative to no motion in that area (Hoffman et al., 2013). The *movement of features* in the dynamic expression benefited recognition of fear and surprise. More specifically, dynamic information facilitated processing of change in the temporal-spatial patterns of the features of each expression and helped to differentiate between fear and surprise (Sato & Yoshikawa, 2004), and, in other studies, judgments about perceived authenticity of smiles (Krumhuber, Manstead, & Kappas, 2007; Rychlowska et al., 2014). Although the dynamic quality of change in the patterns of features was illustrated with dynamic stimuli, the natural quality of the stimuli was often reduced by experimental manipulations. For example, using computerized manipulations to selectively prevent movement in the lower or upper area of a face (Hoffman et al., 2013) reduced the natural quality of the stimuli.

The speed at which an expression unfolds has been shown to influence recognition accuracy, which can vary with different emotions. Kamachi et al. (2001) reported better accuracy to single static images of peak expressions than to dynamic expressions, however they also showed that velocity - the rate of change over time - at which the dynamic expressions unfolded affected recognition. They manipulated velocity by adjusting the number of frames presented to reveal the unfolding of dynamic expressions from neutral to peak in three conditions: short (6 frames, 200 ms), medium, (26 frames, 867 ms) long exposure times (101 frames, 3367 ms). Duration of the exposure time affected intensity ratings but not accurate identification of static expressions. All static expressions were rated as most intense when presented for longer times. In comparison, velocity differentially affected intensity ratings for emotions in dynamic displays. Sad expressions were rated as more intense in longer displays, happy and surprised were more intense in short displays, and anger most intense in the medium display. Of the three velocity conditions, accurate recognition of the sad and angry dynamic presentations was best when the expression unfolded at their medium exposure time (Kamachi et al., 2001). Of particular interest, participants were more accurate recognizing the dynamic compared to static expressions presented in the medium exposure time. Although the researchers acknowledged these differences in values between dynamic and static expressions, they were not analyzed. This point was particularly relevant to the current work because their medium exposure time was most ‘natural’. It was closest to real time (i.e. real time film speed is 24 frames per second) and more importantly, there was no static advantage when the dynamic expressions were closest to natural.

Although there have been inconsistencies with specific outcomes, the behavioral studies consistently indicated differences in underlying systems for the processing of dynamic and static facial expressions and studies with clinical populations have supported such a hypothesis. For instance, Humphreys, Donnelly and Riddoch (1993) reported a case of an agnosic patient who demonstrated poor recognition of static facial expressions but was quite proficient at categorizing dynamic expressions presented in point light displays. Individuals diagnosed on the Autism Spectrum (Gepner et al, 2001; Pelphry & Carter, 2008; Sato, Toichi, Uono, & Kochiyam, 2012; Uono, Sato, Toichi, 2010) and with schizophrenia have also shown better recognition of dynamic compared to static expressions (Kohler, 2003) or differential recognition of static and dynamic displays that varied with positive and negative symptoms of schizophrenia (Johnston et al., 2008). Additionally, individuals with extensive bilateral lesions to the anterior and posterior temporal lobe, and medial frontal cortex showed limited recognition of static faces (e.g. identifying only happiness) yet recognized more emotions in dynamic than static expressions (Adolphs, Tranel, & Damasio, 2003; Atkinson & Adolphs, 2005; Heberlein & Atkinson, 2009). The pattern of findings from clinical studies provided strong support for the notion of separable neural substrates for the processing of static and dynamic expressions.

5.1.i Neuroimaging evidence for separable systems supporting processing of dynamic and static facial expressions

Neuroimaging studies have repeatedly shown that processing a dynamic expression activates *more widespread neural circuitry* than does processing a static expression (Kilts et al., 2003; Sato & Yoshikawa, 2004; Trautmann, et al. 2009, 2013) and that some overlapping areas of activity differ in magnitude (Jing et al., 2014; Furl, Hadj-Bouziane, Liu, Averbeck & Ugerleider, 2012; Pitcher, Dilks, Saxe, Triantafyllou & Kanwisher, 2011;

Sato & Yoshikawa, 2004; Sato, Kochinama, Yashikawa, Naito & Matsumura, 2004; Schultz & Pilz, 2009)[8]. For instance, in contrast to static expressions, processing of dynamic expressions was associated with a greater span of neural activity in cortical and sub-cortical areas associated with perception of changeable features of face processing, motion, biological motion, cognitive and emotion processing, and comprehension of complex mental states (Arsalidou, Morris & Taylor, 2011; Pitcher et al., 2011; Sato & Yokishawa, 2004; Trautmann et al., 2009, 2013). In comparison, processing static expressions has shown less widespread activity that is typically concentrated in specific regions for visual face processing and motor planning (Arsalidou, Morris & Taylor, 2011; Enticott et al., 2008; Kilts et al., 2003; Pitcher et al., 2011).

Neural activity in the visual system associated with face processing has been shown to support processing of both static and dynamic expressions. Within this region, activity in the occipital face areas (OFA)[9] has shown bilateral response to both static and dynamic expressions with no differential response between qualities of expressions (Pitcher et al., 2011). Bilateral activity in the fusiform face area (FFA) is typically sensitive to non-changeable or invariant properties of a face such as face structure and identity (Allison Puce & McCarthy 2000; Haxby et al. 2002; Hoffman and Haxby 2000; Kanwisher, McDermott & Chun, 1997). Levels of activity in the FFA have been recorded in response to both static and dynamic expressions, sometimes with no significant difference in levels between the qualities of expressions (Arsalidou, Morris & Taylor, 2011; Pitcher et al., 2011; Schulz & Pilz, 2009), but more often with greater magnitude of

---

[8.] The imaging studies reported increased levels from baseline during observation of expressions.
[9] Labels used for identifying regions of interest vary across neuroimaging studies. For instance, activity in the inferior occipital gyrus (IOG) and fusiform gyrus (FFG) in response to static faces as reported by Schulz and Pilz 2009 likely correspond to the occipital face area (OFA, Hoffman & Haxby, 2000) and fusiform face areas (FFA, Kanswisher et al., 1997), respectively.

activity in response to static than dynamic expressions (Kilts et al., 2003; La Bar et al., 2003; Sato & Yoshikawa, 2004; Schulz & Pilz, 2009; Trautmann et al, 2009).

Another important component of the visual face processing system is the superior temporal sulcus (STS), of which activity in the posterior (pSTS) and anterior regions (aSTS) is associated with changeable or "dynamic" aspects of a face (i.e. movement of mouth, eyes and *expression*) (Allison Puce & McCarthy 2000; Haxby et al. 2002). Both regions of the STS have shown robust activity bilaterally in response to dynamic facial expressions of emotion (LaBar et al., 2003; Pelphry & Carter, 2008; Schulz & Pilz, 2009) relative to static expressions (Arsalidou, Morris & Taylor, 2011; Pitcher et al., 2011; Sato & Yoshikawa, 2004; Trautmann et al., 2009, 2013).

Activity in the right pSTS and aSTS is not restricted to processing a face. Both areas respond to motion and biological motion depicted in point light displays (Atkinson, 2007; Heberlein & Atkinson, 2009) as well as natural images of implied body motion (Puce et al., 1998) and implied face motion (Jellema & Perrett, 2003; Puce et al., 2003; Schulz & Pilz, 2009).[10] Sensitivity in the pSTS and aSTS to qualities of facial motion coupled with activity in the occipital face areas (aka occipital gyrus) in response to dynamic expressions suggests neural mechanisms that support integration of form and motion during processing of facial expressions (Schultz & Pilz, 2009). Furthermore, selective robust activity in the right aSTS and pSTS in response to perceived dynamic but not static expressions (Pitcher et al., 2011) yet again supported the argument for differences in systems for processing dynamic and static facial expressions.

[10] An example of implied body motion is a static image of an individual swinging a tennis racket, which sets up expectations in the viewer's visual processing for the completion of the action. In comparison, a static image of an individual simply holding a tennis racket does not imply motion. The same principal can be applied to facial expressions; A static image of an expressive face depicts motion in the face (Jellema & Perrett, 2003). It will be proposed in Chapter 6 of this thesis that implied motion in a static image of a face activates the mirror system to complete the action of the expression.

Other areas that have shown different levels of neural activation to static and dynamic facial stimuli are implicated in representation of actions and intentions. Enhanced activity in the parietal cortex, implicated in action understanding and recognition of mental states, has been associated with processing dynamic more so than static expressions (Kilts et al., 2003; Pitcher et al., 2013; Sato et al., 2004; Trautmann et al 2009; 2013). In additional to the inferior-parietal cortex and superior temporal cortex, an important area of the action understanding system identified by Carr et al. (2003) is the frontal region including the medial prefrontal cortex (MPFC), near the superior frontal gryrus, and the inferior frontal gyrus. This frontal region has been associated with processing static and dynamic facial expressions but a greater reliance on the motor regions has been suggested for processing static faces (Kilts, 2003). Kilts et al. (2003) suggested this reliance on the motor regions for processing static faces could be explained by a less distributed area of activation identified for processing static faces - less distributed suggests less dispersion of neural activity supporting processing.

Relative to dynamic expressions, areas of activation associated with processing of static expressions are fewer in number and are primarily connected to the mirror neurons system. In addition to visual areas (i.e. the OFA and FFA reported above), activity in the frontal gyrus has been frequently associated with passive observation of static facial expressions (Kilts et al. 2003; LaBar et al. 2005; Sato, Kochiyama, Yoshikawa, Naito & Matsumura, 2004; Shamay-Tsoory et al. 2009; Shamay-Tsoory 2011; Trautmann et al., 2009, 2013). As reported in Chapter 3, BAs 44, 45, 46 and 6, portions of the frontal gyrus, constitute premotor areas and the more ventral regions along the *pars opercularis* contain motor mirror neurons, an important component of the perception-action coupling system shown to support overt imitation *and* passive observation of facial expressions (Carr et al., 2003; Dapretto et al., 2006; Hennelotter et al., 2004; Iacoboni, 2006; 2008; Jabbi & Keysers, 2008; Vrticka et al., 2013). Reports of activation of these areas during *passive* viewing of static expressions is even more important to the current work than evidence of activation during overt imitation of expressions. Activation during passive observation is indicative of spontaneous, non-intentional, covert motor activity associated with processing a perceived static expression, and provides strong support for the shared-representation model of expression recognition.

The medial prefrontal cortex (MPFC) has been identified as an important cortical area activated in response to emotional cues and supportive of general emotion processing and in particular emotion recognition and regulation (Shamay-Tsoory, 2007; 2011). The resonance mechanism, proposed by Balconi & Barolotti (2012), that supports motor simulation of a perceived facial expression is located in the ventral portion of the MPFC. Stimulation of the MPFC by TMS procedures modulated facial activity *and* emotion recognition accuracy in response to facial expressions (Balconi & Canavesio, 2013). Of particular interest to the current work, data supporting a resonance mechanism were collected in response to observation and imitation of *static* expressions. The research was not conducted on dynamic expressions.

In summary, there are more widespread areas of activation associated with processing dynamic relative to static expressions. Some of these areas overlap with areas associated with processing static expressions and show a greater magnitude of activity to processing dynamic relative to static expressions. In fewer regions that overlap with processing static expressions, greater levels of activity have been, but not always, reported during processing of static relative to dynamic expressions. These areas are in the prefrontal region, particularly the premotor and primary motor areas. High levels of activity in the premotor motor system associated with static expressions led to the hypothesis that the motor system was especially relevant for processing static expressions (Kilts, et al., 2003). As a result, it was suggested for the current work that the motor system was strongly implicated in simulations of static expressions.

Reports of widespread activity associated with processing dynamic facial expressions, including visual, motion, motor, cognitive and emotion systems, led to the suggestion in the current work that simulation of a dynamic expression is extensively multimodal. It has been suggested that due to the abundance of information in a dynamic facial expression, processing a dynamic expression requires top-down, cognitive control (Arsalidou et al., 2011). The reports reviewed above not only supported the argument for differences in processing of static and dynamic expressions (Kilts, 2003; Trautmann, 2009, 2011), they also suggested the hypothesis in the current work that spontaneous facial mimicry would be more vital to the simulation of another person's emotional state following perception of a static expression than following perception of a dynamic expression.

5.1.ii Is there a different function for mimicry in situated expression recognition in response to dynamic and static faces?

Evidence for spontaneous mimicry was generally measured in responses to static images of facial expressions (Dimberg et al. 1998, 2000, 2011). However, there was emerging evidence of mimicry elicited by observation of dynamic expressions. Spontaneous facial activity in response to dynamic expressions has been documented with the Facial Action Coding system (Sato & Yoshikawa, 2007) and facial EMG measures (Lee et al., 2006; Rymarczyk, Biele, Grabowska & Majczynski, 2011; Hess & Blairy, 2001; Sato, Fujimara and Suzuki, 2008, Vrticka, et al., 2013; See also Hess and Fischer, 2013 for a review). As previously discussed, most studies that tested the role of mimicry used static images of expressions or computer morphed animations. Until the current work, no other known studies tested the causal effect of mimicry on situated expression recognition by disrupting mimicry during perception of natural dynamic facial expressions presented with scene context. Furthermore, no other study directly compared the effect of disrupted mimicry on dynamic *and* static presentations of the same expressions in the same study.

5.1.2 Predictions for Experiment 2

The hypothesis for a dynamic advantage in facial expression recognition was tested in this second experiment. Due to the richness of information available in a dynamic facial expression, supported by multiple systems for perception, it was expected that recognition accuracy would be greater to dynamic relative to static expressions.

As outlined above, perceptual processing of information contained in the dynamic expression is associated with multiple, specific areas of neural activity that are specifically associated processing of visual, and motion qualities of information and with emotion, motor, and cognitive responding. Reports of greater intensity of emotion experienced in response to dynamic compared to static facial expressions (Yoshikawa & Sato, 2006) and higher levels of spontaneous facial activity during perception of dynamic, compared to static, expressions, were recorded as EMG measures (Rymarczyk et al., 2011; Sato, Fujimura & Suzuki, 2008), and by a facial action coding system (FACS, Ekman & Friesen, 1976). The results indicated great overlap of systems that support both the perception and production of a dynamic expression. These examples of fine-grained motor activity, plus neuro-imaging evidence for more widespread brain activation underlying dynamic expression, relative to static expressions, supported the hypothesis that an internalized simulation of a dynamic expression would be more robust than simulation of a static expression. According to simulation theorists, activation of multiple systems that support perception of an emotional expression overlap (at least partially) with systems to support experiencing the same emotion. Shared representations for perception and action contribute to the observer's simulated experience of the perceived emotion expression, which is implicitly attributed to the other person (Gallese, 2006; Goldman & Sripada, 2005; Niedenthal et al. 2007; Oberman, et al., 2007; Preston & de Waal, 2002). Thus, a proposed stronger, multi-system simulation of a perceived dynamic expression relative to a static expression led to the prediction of a dynamic advantage in expression recognition in the current work. Although evidence has been found for a dynamic advantage in artificially manipulated dynamic facial stimuli, the evidence for a dynamic advantage in natural dynamic expressions that were presented with context was still lacking.

There was also the question of how context would interact with static and dynamic expressions. The same task for testing situated expression recognition from Experiment 1 was applied in this experiment. Thus, expressions were simultaneously presented with a context scene. As required in Experiment 1, participants simply judged the model's response as either a true or the false response to the context scene.

Consistent with previous research, in this task, processing the context scene was expected to prime perception of the facial expression. The participant's multimodal embodied simulation to the context scene was expected to influence his or her perception and judgments about the model's expression as the appropriate response to the given context scene. As reviewed in the Chapter 2, evidence from past studies has shown that contextual information primes perception of facial expressions. In the task for this experiment, when expressions were presented with context scenes that were congruent in valence, it was challenging to discriminate the true from false response, and it was predicted that participants' embodied response to the context scene would prime their perception of expressions and thereby influence a judgment as '"true". When uncertain, it was hypothesized that the observer would expect the model's response to match his or her own response to the context scene. With this in mind, their embodied simulation of the context scene was expected to prime perception of the facial expression, particularly the difficult-to-discriminate expressions that are presented with a context scene of the same emotion valence. The embodied response would prime perception of the expression as a match to his or her own response to the given context scene of the same valence. A perceived matching state to the same context scene would lead to the judgment of the expression as the 'true' response.

Consistent with Experiment 1, the influence of context on perception and subsequent judgment of the expression was operationalized as the tendency to judge the expression as the true response. The effect would manifest as a greater frequency to judge the model's response to the shared context scene as the "true" rather than false response, resulting in higher levels of accuracy for the true than false response. This effect was observed in Experiment 1 on dynamic expressions. Results from Experiment 1 showed a significant bias for participants to judge the dynamic expression as the true response, especially when the expression and context scene expressed a congruent valence emotion. A replication of such a tendency in the second experiment study was expected to provide additional support for a modulating effect of context on processing emotional information in expressions.

With less information available in a perceived static expression relative to a dynamic expression, and a less robust simulation of a static expression relative to a dynamic expression, it was predicted that perception of a subtle static expression was more vulnerable to a priming influence of the context scene, than perception of the dynamic expression. Thus, it was expected that context would have a *greater* influence on judgments about static that dynamic expressions.

In a previous study, the way in which context interacts with mimicry was explained by congruence. Incongruity between simultaneously self-executed and perceived actions of another individual performing the same task slowed down motor response times (Sebanz, Knoblich & Prinz, 2003). Incongruence between an embodied emotional state and execution of motor tasks with a positive or negative meaning slowed down response times to a perform a motor task (e.g. pushing or pulling joystick either away from or toward self – pushing away is congruent with a negative emotional state, pulling toward self is a congruent with a positive emotional state) (Niedenthal et al., 2007). Additionally, emotional congruence between perceived stimuli, such as context scenes and artificially embedded facial expressions, was rapidly detected, whereas incongruence between the two slowed response times (Righart & de Gelder, 2008). In these examples, congruence between perceptions of two events or between events involving self and another stimulus was rapidly detected and incongruence interfered with motor responses. Thus, it was proposed for the current experiment that congruence between the context scene and the expression of the model would be rapidly detected. More specifically, simulation was the proposed mechanism for detection. The explanation offered here is motivated by findings of Sebanz, Knoblich and Prinz (2003) that complementary actions between one's self and another person are represented in functionally equivalent ways. It is suggested here that representations of two simulated events that are compatible (congruent) would share greater amounts of overlap, whereas, there would be less overlap when representations of two simulated events are incompatible (incongruent). It was hypothesized that with greater overlap in representation, congruence between one's embodied response to the perceived context scene and simulation of the perceived expression of the model would be rapidly detected. Detection of congruence would support formation of a judgment about the expression as the appropriate response – the "true" response - to the given context scene. Detection of incongruence would support a judgment of the expression as the false response. Importantly, incongruence would be more easily detected with equally robust representations that are expected in response to processing dynamic expressions.

How context would differentially interact with mimicry to static and dynamic expressions was uncertain. A simulation of greater magnitude to dynamic than to static expressions is a possible explanation for the lack of effect of disrupted mimicry reported in Experiment 1, which differed from results reported in the literature (Niedenthal, et al., 2000; Oberman et al., 2007; Rychlowska, et al., 2014). A simulation of lesser magnitude for processing static expressions may also help to explain the importance of mimicry for recognition of static more than dynamic expressions. It was suggested above that the multisystem processing of a dynamic facial expressions would generate a multi-system simulation of the expression that is not reliant on facial mimicry. In contrast, with proposed reliance primarily on activation of the motor area for processing static expression (Kilts et al., 2003, Trautmann et al. 2009, 2013) mimicry was identified as an important factor in the simulation of a static expression (Iacoboni; 2006, 2008; Oberman et al., 2007). Mimicry triggers activation of a resonance mechanism (Balconi et al., 2011; 2013) and enhances motor activity, and thereby strengthens the signal from the IFG to activate downstream emotion responses (as per Iacoboni, 2008). For such processing of a static expression, the inclusion of mimicry would likely have a beneficial effect on the simulation of a static expression. A replication of the disruptive effect reported by Oberman et al. (2007) on processing static but not dynamic expressions would corroborate the hypothesis that facial mimicry is an important element in the simulation of static but not dynamic expressions.

If mimicry is an important component in the simulation of a static expression, then there was an expectation that a disruption to mimicry should weaken the simulation of a static expression. In the current study, the robust representation resulting from a multimodal embodied response to a context scene was expected to overwhelm a less robust representation resulting from a less complex simulation of a static expression. There are two considered outcomes for a weakened simulation: 1) rapid detection of incongruence is less likely, and 2) simulation of the facial expression is vulnerable to influence by the embodied response to the context scene. Thus, disrupted mimicry during perception of a static expression would effect a tendency to judge the expression as the true response to the given context scene. This possibility was specifically addressed in the experiment reported in this current chapter, in which an effect of disrupted mimicry on recognition of situated static and dynamic expressions was compared.

Additionally, a differential effect of disruption to mimicry on the processing of positive and negative expression was investigated and reported in this current chapter. When applied by Oberman et al. (2007), the bite-pen manipulation that was intended to disrupt mimicry impaired recognition of *positive* more than negative *static* facial expressions. According to Oberman et al., the “bite” manipulation generated constant activity in the lower facial region, the same region primarily activated in the production of a positive expression. Activity in the upper face region supports the production of negative more than positive expression. During the manipulation, motor activity in that lower region is unavailable for spontaneous mimicry and should affect the observer’s simulation of positive expressions more than negative expressions. Adopting the same rationale in the current work, the labels of “positive” and “negative” denoted emotions that are primarily expressed in the lower and upper areas of the face, respectively.

A direct comparison of how a differential effect of disrupted mimicry on recognition of positive and negative expressions would interact with the dynamic or static quality of natural facial expressions has not been addressed in the literature. Based on the proposed reliance on motor mimicry for successful simulation of a static expression (more so than a dynamic expression) and the unavailability of motor activity required for spontaneous mimicry of a positive expression (caused by the manipulation), it was expected that disruption to mimicry would impair recognition of positive static expressions more than negative static expressions. Secondly, if motor mimicry is less essential to the simulation of a dynamic expression (because of multiple systems that support the simulation of a dynamic stimulus) then disruption to mimicry was not expected to cause a significant impairment in recognition of either positive or negative dynamic expressions. Results from the previous experiment in this thesis did not show impaired recognition accuracy to dynamic expressions caused by disruption to mimicry. Thus, the beneficial effect of mimicry on recognition may be limited to static expressions.

5.2 Method

5.2.i Participants

One hundred individuals (Mean age = 20 yrs, ranging from 16 to 33 yrs, 92% female) volunteered to participate in the current study. All participants were undergraduate students at Dalhousie University, recruited through the Department of Psychology and Neuroscience volunteer subject pool and received 1% credit assigned to one psychology course. Ninety-three participants declared right hand dominance, and all participants declared normal or corrected-to-normal vision. According to the protocol approved by Dalhousie SSHREB, all participants signed an informed consent document.

5.2.ii Materials

For the current study, facial displays and selection of context images were from the same bank of stimuli used in the previous experiment outlined in Chapter 2. The final stimuli consisted of 24 images from the previously selected 40 IAPS images (Lang et al, 1999) and corresponding facial responses to those images by the same four models.

Context stimuli

Consistent with the first experiment, context images were organized into pairs so that each facial response would be presented with the image to which it was the actual response (the "true response") and one foil image (the "false response"). The actual and foil image made one pair. Unlike the first experiment, in the current study the paired context images were always congruent in valence - both intended for either negative or positive arousal. This adjustment to materials was intended to allow for investigation for a differential effect of disruption to mimicry on positive and negative stimuli. Twenty-four context images were sorted into 12 pairs and presented in jpg format. A random order generator determined the pairings of context images.

Face stimuli

Facial responses to the 24 context images by four models (2 male, 2 female) were applied in the current study. The responses were from the same bank of facial stimuli used for Experiment 1. Each response was generated as a dynamic 5 s video in *mov* format *and* a static image in *jpeg* format. The dynamic formats were congruent with the previous experiment – the 5 s dynamic format was a continuous clip of the natural expression. There was no editing of the expression within the 5 s clip.

Static images of the facial expressions were created from the corresponding 5 s dynamic mov files. The dynamic files were imported to iMovie (Apple Inc.). Following the methodology of Kilts et al. (2003), from each of the 5 s dynamic movies a static image was extracted at the perceived apex of the expression. The apex was the point of the response at which the expression peaked. The apex was edited as a single "clip" on iMovie, then exported and saved in jpeg format as a static image with duration of 5 seconds. Static and dynamic facial responses were presented with a resolution of 480 x 640 pixels.

Stimulus organization and presentation

Paired context images were organized into four sets, with each set consisting of three pairs of context images – either two negative and one positive pairs, or one negative and two positive pairs. Context images were presented once with the true response and once with the false response for each of the four models. Images and facial expressions were then sorted into 4 blocks by a Latin Square design. Each block contained all four sets of images but which model's expressions were included varied across the 4 blocks. Thus, participants were exposed to 4 models' facial expressions that were the true and false response to 4 sets of context images across 4 counterbalanced blocks. Order of presentation was randomized within each block

5.2.iii Experimental Design

In the current study, the presentation mode of the facial stimuli was the between-group factor. Participants were exposed to either dynamic *or* static facial expressions. As before, disrupted mimicry by means of a pen manipulation was a within-subject factor. Over alternating blocks, participants experienced half the trials with the pen manipulation and half the trials without the pen manipulation. The experiment conformed to a 2

(Group: dynamic/static faces) x 2 (Disrupted Mimicry: pen/nopen) x 2 (Model Response: true/false) x 2 (Valence: positive/negative) x 2(Pen 1st; pen 1st, pen 2nd) x 4 (Presentation Order, 4 orders of 4 blocks) mixed factor design with Group, Pen Order and Presentation Order as a between-group variables.

There were 24 images, each presented with a true or false response from the four models, and there were 192 trials. Participants did not view every stimulus in both the pen and no-pen conditions. Counter-balancing the 4 sets of stimuli and order of the pen-manipulation (alternating over 1st and 3rd blocks or 2nd and 4th block) ascertained that presentation of a stimulus in pen and no-pen conditions was equally distributed across the sample.

5.2.iv Procedure

Presentation of stimulus corresponded to the previous experiment – a single facial display and context image were presented simultaneously for 5 s, positioned side by side in the centre of an iMac G5 computer screen. Positioning of the image and expression on the left and right sides of the screen alternated with trials.

The procedure followed the basic protocol of the previous experiment but with several modifications. In the current study, there was a single pen manipulation – the "bite-pen" manipulation that had been applied in the previous experiment and by Oberman et al (2007). Participants held the pen laterally between their teeth without the lips touching the pen. This manipulation was intended to disrupt spontaneous mimicry and generate irrelevant feedback "noise" that was meant to interfere with the simulation of the perceived expression.

Participants experienced 10 practice trials followed by 4 blocks of 48 experiment trials with a counterbalanced order of pen and no-pen conditions and presentation order of 4 sets of stimuli. During every trial, participants were instructed to respond “yes” or “no” (by pressing the “z” or “/” keys respectively) to the question “Was the person looking at that image?” and their response was recorded only during the black screen following stimulus offset of the 5 s stimulus.

5.2.v. Data analysis

The data of four participants were dropped due to technical errors in programming. Therefore, data of 96 participants were analyzed.

Primary variables of interest

The dependent variable was the proportion of correct responses (Accuracy). Correct responses were determined by correctly identifying the model’s facial expression as the true or false response.

Participants were quasi-randomly assigned to one of two groups defined by presentation mode of either static or dynamic expressions. Each group defined by static or dynamic facial stimuli had an n of 48. Every participant observed 50% of trials while mimicry was disrupted and 50% while they were free to mimic. The variable “Disrupted Mimicry” identified trials with and without the pen manipulation that was intended to disrupt mimicry.

The variable Model’s Response indicated when the model’s expression was the true or false response to the context image. This variable allowed for a measure of difference in participants’ ability to correctly recognize when the expression was the true response or the false response to the context image. It also provided a measure of an effect of context on judgments of the facial expression. In the previous experiment, participants showed a bias to judge the face as the “true” response when the facial expression was presented with a context image were congruent in valence. In the current experiment, all foil context images were congruent in valence. A replication of the bias would support a hypothesis for a priming influence of context on the observer's judgment of the facial expression.

The variable 'Valence of Stimuli" identified the intended arousal of the context image and facial expression in a given trial as either "positive" or "negative". This variable enabled testing for a differential effect of disrupted mimicry on processing positive and negative expressions that was observed by Oberman and colleagues (2007).

Secondary control variables

Application of the pen manipulation intended to disrupt mimicry was counterbalanced over alternating blocks. Forty-eight participants experienced disrupted mimicry with the pen manipulation throughout the 1^{st} and 3^{rd} blocks ("Pen $1^{st"}$) of trials and 48 participants experienced the pen manipulation throughout the 2^{nd} and 4^{th} blocks ("Pen 2^{nd}"). Order of experiencing the manipulation, identified by a variable labeled "Pen 1^{st}", was added to the main model to determine if disruption to mimicry interacted with the order of experiencing the manipulation.

Presentation order of stimuli was also counterbalanced: There were 4 sets of stimuli presented in one of four possible orders. Thus, another variable to control for order effects was labeled "Presentation Order" and added to the main model.

Analysis of Variance

The dependent variable, the mean proportion of correct responses was calculated for every subject and submitted to a mixed analysis of variance (ANOVA) based on the main

2x2x2x2 model (Group x Disrupted Mimicry x Model's Response x Valence) outlined in section 5.2.iii. Control variables Presentation Order and Pen 1st were added to the model and a preliminary 2x2x2x2x2x2 ANOVA was run. With confirmation of no order effects, both variables were removed from the main model. The *F* values from the main model correlated with the *F* values from the full model with control variables, hence working with the smaller model was not considered problematic.

Due to *a priori* hypotheses of differences in Accuracy due to the static and dynamic quality of facial expression, the data from each group defined by static or dynamic facial stimuli were submitted to separate ANOVAs for more select analyses of the effects of the primary variables. Unless otherwise stated, paired *t* tests or one-way ANOVAs were conducted to test for simple effects. Errors bars represent the standard error of the mean (SEM). Analyses of variance were conducted using R version 2.15, with application of the "*ez* package", version 4.2.2, (Lawrence, 2009).

Alpha, effect sizes, and outliers

The size of an effect was reported as general *eta* squared, which reliably measured the proportion of total variance in a dependent variable explained by the independent variable. Effect sizes calculated for general *eta* were expected to be smaller than the calculation for partial *eta* but are considered a more conservative measure for analysis of multiple factors (Bakeman, 2005). Alpha levels were initially set at a criterion of < .05 for significance. Because of multiple tests in the ANOVA, the risk of a Type 1 error was great. Therefore, a Bonferroni correction was applied. Sixty-three tests were run on the full model containing the control variables, and the corrected alpha was .0008. Once the control variables were dropped from the analysis, there were 15 tests and the corrected alpha was .003.

A criterion for subject outliers was set as any subject mean that was greater than three standard deviations from the group mean. Subject means beyond that criteria were considered an outlier and their data dropped. There were no significant outliers in the accuracy data therefore no data were dropped.

5.3 Results

5.3.i. Control variables

There was no main effect of Order of Presentation of the four blocks of stimuli, $F(3,80) = 4.891$, $p = .99$, ${}_{g}\eta^2 = .00$, nor was there a main effect of Pen 1st, $F(1,80) = 4.905$, $p = .49$, ${}_{g}\eta^2 = .00$. Only four-way interactions were observed involving the variables of Order of Presentation and Pen 1st with Disrupted Mimicry and Valence, $F(3,80) = 15.898$, $p < .0008$, ${}_{g}\eta^2 = .02$, and with Disrupted Mimicry and Model Response, $F(3,80) = 12.251$, $p < .0008$, ${}_{g}\eta^2 = .03$. Further testing revealed the interaction of Order of Presentation, Pen 1st, Disrupted Mimicry and Valence was significant on Accuracy when expressions were static, $p < .0008$, ${}_{g}\eta^2 = .02$, but not dynamic, $p > .0008$, ${}_{g}\eta^2 = .02$. Furthermore, the interaction of Order of Presentation, Pen 1st, Disrupted Mimicry and Model Response was significant when expressions were dynamic, $p < .0008$, ${}_{g}\eta^2$ .07, but not static, $p > .0008$, ${}_{g}\eta^2 = .01$. Further examination of the means did not reveal any systematic effect of Order of Presentation or Pen 1st within these interactions. Consequently, they are not reported in any further detail.

5.3.ii General findings

The total mean rate of accuracy was 60%($SD = .03$, $N = 96$). This rate was lower than the previous experiment (70% /.05), which was expected as a consequence of including only the more difficult same-valence trials. Accuracy was, however, higher than comparable studies (e.g. 52% accuracy on tasks measuring recognition of complex mental states, Back et al., 2009). In the current study, no participant means for accuracy were less than 50%. The mean rate of 60% was significantly different from chance, t (190) = 32.6599, p .0001, *Mean Difference* = .10, *SED* = .306, *CI* = 9.3960 - 10.6040.

The ANOVA table for the main model, and for select ANOVA on data by the Dynamic and Static groups, are included in the Appendix. With the corrected alpha level of .003, a main effect of Group, defined by static and dynamic expressions, was not significant on Accuracy, $F(1,94) = 8.692$, $p = .004$, ${}_g\eta^2 = .005$. The means indicated participants were accurate when expressions were dynamic compared to static, but the differences were not significant, $M_{\text{Dynamic}} = .61$, $SD = .03$, $M_{\text{Static}} = .59$, $SD = .035$. This initial finding did not support the hypothesis of better Accuracy when expressions were dynamic compared to static.

Model Response was identified *a priori* as an important factor for understanding the effect of context on recognition. An influence of context was operationalized as higher levels of judging the model's response as true *vs* false. Collapsed across data for static and dynamic expressions, higher levels of accuracy to the true than false response were observed and explained a significant proportion of variance in accuracy, $F(1,94) =$ 17.451 $p = .00007$, ${}_g\eta^2 = .08$.

Because of *a priori* expectations for a difference in processing dynamic and static expressions and a hypothesis of greater effects of context on static relative to dynamic expressions, select analyses were run on data from groups who saw dynamic and static expressions. Select ANOVAs revealed a main effect of Model Response on accuracy when expressions were dynamic but not static. Participants who saw dynamic expressions were significantly more accurate recognizing the Model's Response as true than false, $F_{Dynamic}(1,47) = 19.067$, $p = .00007$, ${}_{g}\eta^2 = .15$. Judging the response as true or false accounted for 15% of the total variance in accuracy to dynamic expressions. In comparison, when expressions were static participants were also more accurate recognizing the true than false response but the difference in accuracy was not significant, $F_{Static}(1,47) = 3.342$, $p = .07$, ${}_{g}\eta^2 = .04$. Levels of accuracy for recognition of the True response, as shown in Table 5.2, were significantly lower when expressions were static compared to dynamic, ($p = .02$, ${}_{g}\eta^2 = .05$). Hence, Model Response, as an index of context, affected judgments about dynamic but not static expressions.

Table 5.1 Mean rates of accuracy for recognition of true and false responses when expressions were dynamic ($n = 48$) and static ($n = 48$)

	Mean Accuracy / SD	
	TRUE	FALSE
Dynamic	.67/09*^	.56/09^
Static	.62/10*	.57/12

*^Significant @ .05

Judging the model's expression as the true rather than false response - as a match to the participants' response to the given context scene - was a relevant factor affecting the recognition process of dynamic expressions and supported the hypothesis for an influence of context on processing the expression. That the effect was stronger on recognition of dynamic than static expressions was unexpected.

There was not a main effect of disrupted mimicry, $F(1,94) = 1.339$, $p = .25$, ${}_g\eta^2 = .00$. There was no difference in accuracy when mimicry was disrupted and when participants were free to mimic, regardless of whether expressions were static, $p = .43$, or dynamic, $p = .40$.

It was hypothesized that Disrupted Mimicry would impair recognition to positive more so than negative expressions. The interaction of Valence with Disrupted Mimicry was not significant, $p = .25$, when expressions were dynamic, $p = .31$, or static, $p = .89$. Participants were always more accurate when facial expressions and context scenes were positive than negative, $F(1,94) = 130.176$, $p < .0001$, ${}_g\eta^2 = .055$, and Valence did not interact with Group, $p = .94$. Better recognition of positive than negative expressions was observed when expressions were static and dynamic, and did not change as a function of mimicry. As shown in Table 5.2, contrary to expectations, disrupted mimicry did not have a greater effect on recognition of positive than negative expressions, when expressions were dynamic or static. These results were not consistent with impaired recognition accuracy to static happy expressions caused by the same manipulation, as reported by Oberman et al. (2007). The means in Table 5.2 indicate better recognition of negative dynamic expressions during the disrupted mimicry compared to the free mimicry condition, but the difference was not significant.

Table 5.2 Mean rates of accuracy when expressions and scenes were positive or negative did not change as a function of mimicry

		Accuracy Mean /SD	
		Mimicry	
		Disrupted	Free
Dynamic	Positive	.65/.07	.65/.05
	Negative	.60/.06	.57/.07
Static	Positive	.63/.07	.62/.07
	Negative	.56/.07	.56/.06

It was also suggested that, if mimicry underlies recognition, one role for mimicry is to enable detection of fine-grained details of an expression supporting recognition of a false response. If so, then with such fine-grained detection the influence of context should be less apparent when participants were free to mimic, resulting in no significant difference in accurate recognition when the response was true and false. This was not observed. As illustrated in Figure 5.1, there was no change in higher levels of accuracy for recognition of the true than false response as a function of Disrupted Mimicry, when expressions were dynamic or static.

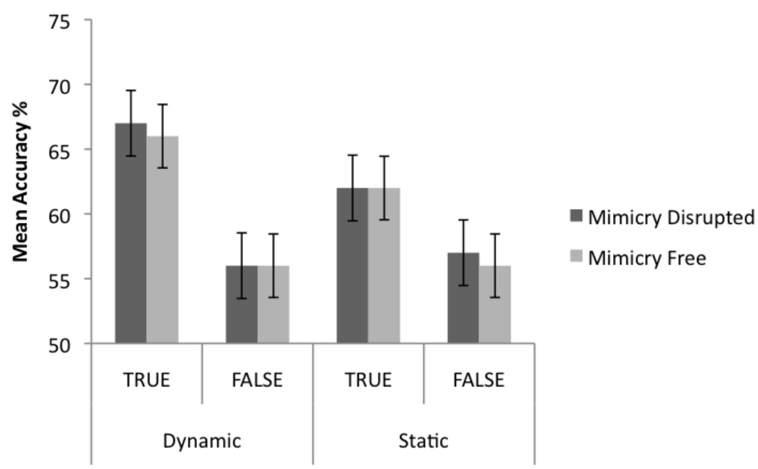

Figure 5.1 The effect of Model Response in dynamic and static expressions did not change as a function of Disrupted Mimicry. *N* = 96, Error bars indicate *SEM*.

5.3.iii Summary of results for processing dynamic and static expressions

Expression recognition was not more accurate when expressions were dynamic than static. The hypothesis for a dynamic advantage in recognition was not supported. There was a robust effect of Model Response on accurate recognition of dynamic but not static expressions. The model's response was more accurately recognized as true than false. Participants showed a bias to judge the dynamic response as true, confirmed by criterion c analysis. Higher levels of accuracy were an index of an influence of context on recognition. Hence the hypothesis of an influence of context on judgments about expressions was supported.

One hypothesized role of mimicry was to reduce the influence of context on recognition of expressions. There was no change in higher levels of accuracy for the true than false response when mimicry was possible. There was no evidence for any impairment to recognition of either dynamic or static expressions when mimicry was disrupted. Nor was there any impairment of recognition of positive expressions more so than negative expressions when mimicry was disrupted. The findings did not support the hypotheses of mimicry underlying situated expression recognition.

5.4 Discussion

The inclusion of static as well as dynamic facial expressions in Experiment 2 was intended to support or not the hypothesis for a dynamic advantage in expression recognition. Moreover, additional objectives were to identify differential effects of context and mimicry on processing dynamic and static facial expressions.

The results from Experiment 2 did not support the hypothesis for a dynamic advantage in expression recognition. Higher levels of accuracy when expressions were dynamic than static were indicated by the means, but the differences did not reach significance. A sector of the literature on expression recognition addressed and compared the validity of testing expression recognition with dynamic or static facial expressions. Reports were conflicted with regards to a dynamic advantage to recognition accuracy when information about an advantage for *natural* dynamic facial expressions was limited. The richness of information found in a dynamic expression that is not found in static expressions has been identified in previous research as an explanatory factor for a dynamic advantage in expression recognition (Ambadar et al., 2005; Bassili, 1978, 1979; Back, Jordan and Thomas, 2009; Hoffman et al., 2013; LaBar et al., 2003; Weyers, Muhlberger, Hefle & Pauli, 2007). Perceptual processing of information contained in the dynamic expression is supported by multiple systems e.g. visual, motor, motion, emotion, cognitive (Trautmann et al, 2009, 2011), Findings for greater intensity of emotion experienced (Yoshikawa & Sato, 2006), higher levels of spontaneous facial activity, (Rymarczyk et al., 2011; Sato, Fujimura & Suzuki, 2008), and widespread brain activation underlying perception of dynamic expression, (Kilts et al., 2003; Trautmann, 2009; 201, 2013) were indicators of a significant overlap in representation for perception and production of naturalistic dynamic expressions. These findings led to the hypothesis of a more robust simulation of a dynamic expression relative to a static expression, which in turn supported the predictions for a dynamic advantage in expression recognition.

Unlike static expressions, the dynamic expressions unfolded over time; they included updated, changing information revealed in the temporal-spatial aspects of the expression as it evolved over the 5 s display. Sometimes the dynamic expression was bimodal, revealing more than one peak or more than one expression in the 5 seconds of changing information. In comparison, the static images displayed one frame taken from the peak of the respective dynamic expression for an equal display time of 5 s and contained no updated or changing information about the expression. Consequently, judgments about a static expression were likely formed prior to offset of the stimulus whereas judgments about the dynamic expression were likely formed close to or following stimulus offset. The longer duration was expected to benefit processing of dynamic expressions. While the means indicated better recognition of dynamic than static expressions, the difference was not significant.

Another objective of the experiment was to determine if the inclusion of context scenes in the task would differentially affect recognition of static and dynamic expressions. As hypothesized, during a social interaction, one person unconsciously expects another person's response to a given, shared situation to be similar to his or her own response to the situation. Moreover, when expressions are difficult to discriminate, possibly because of subtle or blended emotions, the observing individual searches and expects to perceive evidence of a similar, matching response in the other person's facial expression.

As predicted, results showed an influence of context, primarily on dynamic and only minimally on static expressions. The size of the effect was greater on processing dynamic expressions. A response bias to judge dynamic expressions as the "true" response was confirmed. The bias supported the hypothesis that participants' perceptions of expressions were influenced by his or her embodied response to context scenes. An influence of context on perceptual processing of static facial expressions was expected based on previous reports in the literature. That the effect was stronger on the perception of dynamic than static expressions was unexpected.

One explanation for the results may be offered by inverting the rationale that was proposed for static expressions. It was proposed that a weak representation of the simulated static facial expression would be vulnerable to a robust representation of an embodied simulated response to the context scene. Conversely, the more expansive representation generated by a multi-system simulation of a perceived dynamic expression may explain the effect. Multiple systems that support simulation of the perceived complex dynamic expression overlap with multiple systems activated during the embodied response for processing the perceived context scene. The “dual demand” increased the magnitude of activity of visual, cognitive, sensorimotor, and emotion systems. The expansive perceptual processing of the context scene overlaps with, and largely maps onto, the perceptual processing of the dynamic expression. Input from the two sources resulted in even more widespread and greater magnitude of activity to support perception and simulation. The result was a more unified overlap in representation of the perceived and simulated dynamic expression and the perceived and simulated response to the context scene. The overlap supported the integration of the two representations. A greater influence of context was the result of a “better fit” in overlap of represented events. This sequence of processing would also explain the longer response times to dynamic expressions.

The challenge of the task was to discriminate fine details of a natural, spontaneous expression in order to recognize it as the true or false response to the given context scene. In Experiment 2, the models’ true and false responses to a context scene were always of the same valence. Thus, the true and false responses to the paired actual or foil context scenes were both of the same or similar emotion and therefore considered ‘difficult to discriminate’. It had been proposed that mimicry, even covert mimicry of the model’s expression, would support the challenging discrimination process. Mimicry was expected to strengthen the simulation process and thereby strengthen the representation of the facial expression and support detection of either congruence or incongruence between representations of the facial expression and the participant’s response to the context scene. Without mimicry, the discrimination process would be even more difficult. The findings in Experiment 2 did not support the hypothesis that spontaneous facial mimicry modulates the effect of context on recognition of facial natural expressions. There was no change in higher levels of judging the expression as the true response when participants were free to mimic or when mimicry was disrupted. There was no evidence for an underlying role of facial mimicry in reducing the influence of context on recognition of dynamic or static expressions.

A prevalent hypothesis in the literature is simulation of the perceived expression of another individual supports recognition of that expression. Moreover, mimicry was proposed as an important mechanism contributing to the simulation of the perceived expression by increasing activation of the motor system (Iacoboni, 2008). According to this hypothesis, participants more accurately recognize facial expressions when they are free to mimic the expression than when mimicry is disrupted. Disrupted mimicry during the task was expected to interfere with the participant's simulation of the perceived emotional state of the model, and thereby interfere with the recognition process.

With a proposed reliance primarily on the premotor and motor system for processing and simulating static expressions (Kilts et al., 2003; Trautmann, 2011, 2013), it was further hypothesized in Experiment 2 that facial mimicry would support recognition of static expressions in particular, and that disruption of mimicry would have a significant effect on recognition of static expressions. Disruption of mimicry was not expected to be particularly detrimental to the simulation of a dynamic expression because there are so many more systems activated than the motor system. Therefore, it was hypothesized that mimicry is a less vital mechanism for the simulation of a dynamic expression and thereby recognition of dynamic expressions.

A disruption to mimicry did not affect recognition to dynamic or static expressions. The impairment to recognition of static expressions caused by the same manipulation intended to disrupt mimicry reported by Oberman et al. (2007) was not replicated in Experiment 2. It was predicted that the manipulation blocks spontaneous activity of lower facial muscles, particularly the zygomaticus major, typically activated to support positive expressions. Irrelevant feedback caused by the manipulation was expected to cause interference in the neural signaling to the motor system, thus interfering with simulation expressions. Any disruption caused by the manipulation in the current study was not great enough to alter explicit judgments about the expression.

The findings in the current study may be explained by the degree of motor activity necessary for the simulation process. A covert motor response has been shown to sufficiently support perception of features in the static image of an expression (Likowski et al., 2012). EMG measures of facial activity not visually perceived, nor consciously felt, have been associated with accurate recognition of perceived facial expressions (Dimberg et al., 2012). It is known that *both* the primary motor and pre-motor regions are active during imitation and passive observation of static facial expressions (Balconi, 2011; 2103; Carr et al., 2003; Likowski et al., 2012) and so the suggestion raised is that premotor activity might sufficiently support recognition without overt activity in the primary motor areas. Levels of activity are less in the primary than pre-motor system during observation than overt imitation, and passive observation entails spontaneous, low levels of covert motor activity (measured by EEG recordings, Carr et al., 2003 plus EMG and fMRI measures, Likowski et al., 2012). Although *spontaneou*s overt motor activity was prevented it may be that the manipulation did not interfere with the premotor system that would typically support covert mimicry during passive observation of an expression, at least to the extent of causing impairment to recognition.

5.4i Summary

That the findings from Experiment 1 and 2 differ from reports in the literature is not soley explained by the presentation of dynamic and not static expressions. The inclusion of context and the duration of expressions at 5000 ms (as opposed 500 ms presented by

Oberman et al., 2007) need to be considered along with any differences in dynamic and static expressions.

One report from the literature comparing recognition of dynamic and static expressions suggested that a static image of an expression at its fullest requires no additional information for recognition that might be available in the dynamic expression (Bould & Morris, 2008). Yet, other reports identified qualities of information that were not available in the static expression and would contribute to better recognition (Ambadar et al., 2005). Questions arose about “quantity” as well as quality of information that is available in dynamic and static expressions. For instance, what amount of information was sufficient to measure differences in processing static and dynamic expressions? Motion had been identified as one quality of a dynamic expression that can be only implied in a static expression. A number of studies have tested velocity as the important factor affecting the perception of a dynamic facial expression and one that would explain a dynamic advantage in emotion recognition accuracy. Many past studies tested the effect of velocity by artificially manipulating the content of an expression. With the continued goal of increasing ecological validity, velocity was not manipulated in the research for this thesis. Instead, investigations of the dynamic advantage continued by manipulating the amount of dynamic information, such as motion, available in naturally unfolding expressions.

CHAPTER 6: EXPERIMENT 3

The research for Experiment 3 continued to test the theory that recognition of another person's facial expression is supported by an observer's simulation of the perceived expressed state and that the observer's response to a shared context has an influence on judgments about the expression. The research question in this final study was extended to address whether perceived motion cues are a critical feature of dynamic expressions that explain the advantage in recognition of dynamic compared to static expression. Furthermore, it followed to question if the lack of motion in static expressions explains the hypothesized greater reliance on mimicry for recognition of static compared to dynamic expressions.

6.1 Do motion cues underlie the dynamic advantage?

As reviewed in the previous chapter, the proposed dynamic advantage in facial expression recognition had been explained by the rich complexity of information available in dynamic but not static expressions (Ambadar et al., 2005). It was suggested that motion cues contained in dynamic expressions capture attention and increase arousal in the observer (Trautmann, et al., 2009). As outlined in Chapter 5, Ambadar et al. (2005) reported that participants were more accurate in response to dynamic expressions relative to single static and multi-static expressions. Moreover, according to Ambadar et al., perceived change is the key component in a dynamic expression that explains a dynamic advantage in expression recognition accuracy. In a 4^{th} condition of their 2005 study, Ambadar et al. measured recognition accuracy to dynamic and static expressions as well as presentations of the first and last images of an expression (the first frame of the start of an expression and the last frame at the apex of the expression). Recognition accuracy was better in both dynamic and "first-last" conditions relative to single static or multi-static presentations, and there was no difference in accuracy between the dynamic and first-last conditions. It was concluded that better accuracy was not due to perception of unique temporal qualities specific to an expression, because this information was not available in the first-last condition. Instead, better recognition was explained by a particular, beneficial effect of motion that enabled perception of change in the expressions. Their dynamic expressions had been built from multiple sequences extracted from different points of intensity of an expression. Like the dynamic expressions (but unlike the multi-static images), the first and last images were presented without visual noise masking so that there was no disruption to the perception of motion. Ambadar et al. (2005) suggested that motion was inherent to the dynamic quality of an expression, increasing sensitivity to changes in the facial features as the expression unfolds.[11]

Manipulations of speed have been shown to affect accuracy (Ambadar et al., 2005; Bould, Morris & Wink, 2008; Kamachi et al, 2001), higher levels of observers' spontaneous facial activity in response to dynamic than static expressions (Achaibou et al., 2008; Ryzmaryk et al., 2011; Sato & Yoshikawa, 2004; Sato et al., 2008; Weyers, 2006), as well as judgments about the authenticity of dynamic expressions (Krumhuber & Kappas, 2005; Krumber, Manstead & Kappas, 2007; Maringer et al.,2009; Rychlowska et al., 2014). For instance, smiles with longer onset and offset durations[12] were judged

---

[11] Bould, Morris, & Wink (2008) contested claims made by Ambadar et al. (2005), however their interpretation of Ambadar's conclusion was questionable. In their paper, they wrote that Ambadar claimed "recognition was facilitated by perception of change rather than perception of motion", yet when reviewing the same report by Ambadar et al (2005) for this research, special attention was paid to the specifics of their findings and in particular their conclusion that "motion was inherent to the perception of change in a dynamic expression". Hence, the position taken for this research was that Ambadar et al. (2005) were not suggesting that motion was not important in expression recognition, but instead were positing the opposite.

[12] In this review, the terms offset and onset are used as defined by Krumhuber et al who measured the effects of onset and offset durations on judgments of the meaning of smiles. They defined onset duration as the "length of time from the start of a smile until its maximum intensity" and offset duration as "the length of time from the end of the apex until the smile disappears" (2007).

as more genuine than smiles with shorter onset and offsets (Krumhuber & Kappas, 2005; Krumhuber, Manstead & Kappas, 2007).

In addition to the pivotal research by Ambadar et al., 2005, the impact of motion on expression recognition has been studied by manipulating velocity (the speed at which an expression unfolds) and perceived intensity of an expression. Frequently, this was accomplished by morphing techniques to create a dynamic stimulus from neutral and intense (intensely expressive) static expressions. Typically, speed and/or intensity have been manipulated by varying the number of components in a morphed sequence between the neutral and peak expression. This technique allowed for control over the speed of the expression and amount of perceived visual information. Perceived speed of an expression was also manipulated by varying the duration of expressions. For example, Kamachi et al (2001) presented fewer frames over a brief duration to generate a high-speed expression, whereas more frames (or multiple increments) presented in a longer duration generated a slow speed expression. By varying the number of frames from different stages of an expression that were presented over different durations, Kamachi et al. (2001) generated variation in perceived velocity in the unfolding of an expression. Recognition of dynamic sad and angry expressions was most accurate when velocity was closest to real time speed (26 fps/867) compared to when the expressions unfolded over their manipulated fast or low speeds.

6.1.i Duration of dynamic and static expressions and evidence of spontaneous mimicry

Of particular interest are the duration of expressions reported for several behavioral studies that compared levels of spontaneous facial mimicry in response to static and dynamic expressions (some of these behavioral studies included EMG measures). Levels of spontaneous mimicry in response to dynamic expressions were reported for durations of 1520 ms (Sato & Yoshikawa, 2004), 1500ms (Ryzmaryk, 2011; Sato et al., 2008) and 1400ms, (Achiabou et al, 2008) and in some cases, the mimicry to dynamic expressions was greater than to static expressions (Sato & Yoshikawa, 2004; Sato et al., 2008). Hess & Blairy (2001) reported evidence of facial mimicry during perception of 15-second videos containing dynamic expressions. Several studies presented dynamic expressions for an unlimited time, until participants responded (e.g. Ambadar et al., 2005; Maringer et al., 2009; Niedenthal et al., 2001).

In contrast, evidence of facial mimicry to static expressions was frequently measured in response to duration times that were typically brief (e.g. 50 ms, Dimberg, 2000; Stel & van Kippenberg, 2008). Overt manipulation of duration times of static expressions affected the likelihood of spontaneous mimicry to perceived static expressions (Sonnby-Borgstrom, 2002). For instance, evidence of facial activity was measured by EMG activity when durations of static expressions were less than 100 ms but not when duration times were 100-1000 ms (Sonnby-Borgstrom, 2002).

Duration of expression and occurrence of spontaneous mimicry reported in previous studies were of particular interest to the current work because of what they suggest about underlying processes for recognition of dynamic and static expressions. Spontaneous mimicry to dynamic expressions was measured in response to long durations (~1500 ms). Spontaneous mimicry to static expressions occurred in response to brief (<100 ms) as well as long (1500 ms) durations (Sato & Yokishikawa, 2004; Sato et al., 2008). Several researchers reported greater levels of facial activity in response to dynamic than static long expressions. Interestingly, Sato & Yoshikawa (2004) noted that in response to expressions of 1520 ms overall frequency of mimicry was low, at only 20%, but of that 20% there was more mimicry to dynamic than static expressions. There was evidence of mimicry to long dynamic expressions, but unlike responses to static expressions, there was no evidence to suggest that mimicry occurs in response to brief dynamic expressions. Moreover, there was nothing in the literature to indicate that a dynamic advantage would be observed if dynamic expressions were brief.

Implications from the findings reported above are consistent with reports of more expansive cognitive and neural systems underlying the processing of dynamic expressions, relative to static expressions. Motion cues that differentiate the dynamic expression are perceived during early visual processing of a facial expression. The right aSTS and pSTS are associated with changeable aspects of a face such as movement of eyes and mouth (Allison, Puce & McCarthy, 2000; Haxby et al 2002). Increases in activity of these two STS regions were reported during perceived dynamic facial expressions (Labar et al., 2003; Pelphry et al., 2007; Schulz & Pilz, 2009) with greater levels of activity relative to static expressions (Pitcher et al., 2011; Sato & Yoshikawa, 2004; Trautmann et al., 2009, 2013). These findings had contributed to the argument for separable systems supporting the processing of static and dynamic expressions. Moreover, motion cues enhance the saliency of a dynamic expression to capture the observer's attention (Trautmann et al., 2009) and, as established, enable the perception of change in an expression. Both saliency and perception of change have been identified as key components of a dynamic expression that account for better recognition accuracy (Ambadar et al., 2005; Kamachi, et al, 2001; Trautmann et al., 2009). Based on these observations, it would seem that fewer motion cues to capture attention and indicate change in the temporal spatial relations of the facial features, could reduce the saliency and perception of change, and ultimately the dynamic aspect of an expression. Ambadar et al. (2005), proposed perception of change is key to the dynamic advantage. The reported importance of motion led to the suggestion in the current work that without motion in an expression, the dynamic advantage in recognition accuracy would be lost.

The proposed importance of motion to the dynamic advantage led to questioning the effect of motion in processing dynamic expressions. What amount of motion cues in a dynamic expression would make a significant difference to recognition accuracy? It was posited that substantially fewer motion cues, in addition to affecting less saliency and perceived change, would influence a weaker representation of the perceived expression in the observer. According to simulation theory, the observer's simulation of the perceived expression is the process by which the representation of the expression is generated (Gallese, 2003). To that end, the observer's expansive and multimodal simulation typically activated in response to a perceived long dynamic expression containing multiple motion cues would not be experienced to the same magnitude in response to a brief expression containing significantly fewer motion cues. Instead, the observer's simulation of the expression would be reduced and the resulting representation of a brief dynamic expression could be drastically diminished. Hence, fewer motion cues perceived in a brief dynamic expression could be costly, and likely provide insufficient information about perception of change to result in a dynamic advantage in recognition. If perceived motion cues are an important component of a dynamic expression that accounts for better recognition accuracy, then insufficient motion cues should result in levels of recognition accuracy reduced to that of static expressions.

As addressed in Chapter 5, reports from neuroimaging studies suggest that processing static expressions relies primarily on the motor system. One of the hypothesized benefits of facial mimicry to recognition of static facial expressions is increased activation of the motor system enhancing the simulation process and ultimately influencing a richer representation of the perceived expression (Gallese & Sinigaglia, 2008). Importantly, in addition to increased levels of activity in the motor mirror neurons that occur during perception of an expression, there are also increased levels of overt motor production occurring with externalized facial mimicry activity. While there is no perceived motion in a static facial expression, mimicry generates movement of the facial muscles. Facial activity contributes to the enhanced motor activation identified as an important component to the simulation process of a perceived static expression. A weak effect of disrupted mimicry in Chapter 5 provided some support this hypothesis. A disruption to facial mimicry slowed down processing of positive static expressions.

What has not been addressed in the literature, but considered here, is the possible benefit of mimicry to the recognition of brief dynamic expressions. A significant reduction in perceived motion cues in a brief dynamic expression may influence a need to generate produced motor activity that would assist processing. Essentially, a gap in the literature addressing differences between brief dynamic and static expressions led to questioning what might change in the processing of a brief dynamic expression that would be different from processing a long dynamic expression but similar to processing a static expression.

6.1.ii Experiment 3

The difference between static and dynamic expressions of brief and long duration was investigated by the experiment reported in this chapter. An investigation into the effect of context on recognition of static and dynamic expressions continued. Mimicry as an underlying mechanism of emotion recognition to static and dynamic expressions was again tested by disrupting mimicry.

With less dynamic information available in a brief dynamic expression than what is available in a long dynamic expression, reduced accuracy was expected. To what extent the brief dynamic expression would be more advantageous to recognition than a static expression was uncertain. To address these questions and predictions, duration times of dynamic and static expressions were modified in the study reported in this chapter.

Maintaining the goal of increasing ecological validity, investigations for this thesis continued with naturally unfolding expressions. Perceived change in the dynamic expression would be manipulated in the current study by lengthening or shortening the duration of the natural expression. Expressions would be presented in brief or long durations, thereby naturally revealing more of less of the dynamic quality of an expression without artificially manipulating the speed of the evolution of the expression. Unedited sequences from the original 5 s expressions were selected (a single sequence without edits) for brief and long durations. The 5 s expressions had been presented at real time (25 frames / second) and so appeared to naturally unfold. Duration of the expressions was centred on the apex, so that the peak of the expression was perceived in both brief and long dynamic expressions. The velocity – or speed – of expressions was not manipulated. Instead, an unedited sequence of different durations allowed for an investigation into the benefit of motion in naturally unfolding dynamic expressions. Static expressions were a single frame from the apex of the expression and presented for the same long and brief durations.

Only more recently had Rychlowska et al. (2014) presented naturalistic dynamic expressions to test judgments of authenticity of smiles. Otherwise, in the studies reviewed above, dynamic and static expressions were not natural, and dynamic expressions were computer generated, either by morphing techniques with static expressions (e.g. Sato et al, 2008) or generating synthetic facial expressions (e.g. Krumhaber et al., 2011). Additionally, no study had manipulated the unfolding of dynamic expression in a non-artificial way. The speed of change in expressions had always been artificially manipulated. Relatively close to natural time was the medium duration of stimuli (26 frames / 867 ms) presented by Kamachi et al. (2001), to which participants were most accurate. So too were the dynamic sequences presented at 25 frames per 1000 ms by Bould et al, (2008), but their last image of the sequence was held for an additional 1000ms, which compromised the 'natural' unfolding of their dynamic expressions.

Not only were the quality of facial expressions not natural, naturalistic dynamic and static facial expressions had not been presented to measure recognition of expressions that were situated. Context had not been included in the studies measuring the impact of motion on emotion recognition of static and dynamic expressions. Effects on context had been observed on brief static expressions (Aviezer et al., 2008; Hassin et al., 2013; Righart & de Gelder, 2006). There was no information on how context would interact with dynamic expressions of different durations.

6.1.iii Summary & Predictions

Three research questions were addressed in the experiment reported in this chapter: 1) Does motion explain the dynamic advantage in recognition accuracy? 2) Is there an effect of context on emotion recognition to dynamic and static expressions? and 3) Does mimicry underlie recognition to dynamic and static expressions?

With regards to the first question, if perception of motion in a dynamic expression explains the dynamic advantage in recognition accuracy, then a significantly reduced amount of perceived motion should result in decreased levels of recognition accuracy. With brief duration of expressions, levels of recognition accuracy to dynamic expressions may be reduced to the level of accuracy to static expressions.

Relative to long durations (1000 ms), brief durations of the same dynamic expressions (250 ms) were expected to contain fewer changes in temporal spatial relations of the facial features. The difference in perceived motion cues in long and brief durations of dynamic expressions was expected to result in greater accuracy but longer reaction times for long relative to brief duration expressions. More information in the long expression that is beneficial to recognition, such as perception of change over time, would require longer processing time.

If even a reduced amount of motion perceived in brief dynamic expressions (250 ms) is beneficial to expression recognition, then accuracy was expected to be greater to brief dynamic expressions than to brief and long static expressions. If brief dynamic expressions contain sufficient motion cues to at least capture attention, then reaction times should be longer to brief dynamic than static expressions.

Differences in processing static expressions of long and brief durations were not expected since there would be no change in information contained in the expressions. Unlike a longer presentation of a dynamic expression, a longer presentation of a static expression would simply provide the observer with more time to gaze at an unchanging expression. Hence, neither accuracy nor response time was expected to differ with duration of static expressions.

Expressions were presented in this next experiment for brief or long durations, thereby naturally revealing more or less perceived change in dynamic facial expressions. A manipulation of duration of the expressions allowed for an investigation into the benefit of actual motion in dynamic expressions. No difference in recognition accuracy to long and brief durations of dynamic expressions was expected to weaken the argument for motion as a critical factor in the dynamic advantage hypothesis.

The second research question addressed context. A strong effect of scene context on recognition accuracy was observed in Experiments 1 and 2 and a replication was expected in the current study. The influence of context was observed on recognition of both dynamic and static expressions, but more robustly on dynamic expressions. It was expected that context would influence processing of both static and dynamic expressions and the effect of context would be greater on long dynamic expressions than brief dynamic expressions and static expressions. Working from the proposed explanation for the influence of context on dynamic expressions from Chapter 5, less dynamic information available in the brief dynamic expression would allow for less overlap in representations generated by the simulated model's expression and the observer's own embodied response to the context scene.

Consistent with the literature, but inconsistent with findings of Experiment 2, if mimicry does underlie recognition then mimicry was expected to benefit recognition of static expressions, of both brief and long durations. Without perceived motion cues the increased motor activity due to facial mimicry was expected to improve the simulation process of the perceived static expression. If processing brief dynamic expressions is benefited by mimicry, then disruption to motor mimicry was expected to result in lower recognition accuracy and slower response times. Perceived motion cues in a brief dynamic expression may be substantially reduced so that accuracy is also significantly reduced. In that case, then like static expressions, mimicry of the expression may supplement the lack of perceived motor activity. If so, then disrupted mimicry should result in reduced accuracy relative to when participants are free to mimic a brief dynamic expression.

6.2 Method

6.2.i Participants

A total of 65 participants (M_{Age} = 20 yrs, ranging from 18 to 39 yrs) volunteered to participate in the current study. Of these participants 76% identified as female, 92% declared right hand dominance, and all participants declared normal or corrected-to-normal vision. All participants were undergraduate students at Dalhousie University, recruited through the Department of Psychology and Neuroscience volunteer subject pool and received 1% credit assigned to one psychology course. According to the protocol approved by Dalhousie SSHREB, all participants signed an informed consent document.

6.2.ii Adjustments to materials and procedure

The findings of Experiments 1 and 2 influenced decisions to modify testing materials and procedure in the current experiment. Low rates of accuracy due to presentation of only same-valence context pairs were a concern. The decision to use same valence stimuli was based on the findings from Experiment 1 in that the important results about the influence of context had been found in the same valence condition. The results from Experiment 2 were similar, showing a strong effect of context on recognition of dynamic expressions.

Findings from Experiment 2 were inconsistent with the literature and the results often weak. Consequently, in the current study, adjustments were made to the testing procedure and stimuli with the intention of obtaining results that were potentially more reliable regarding the involvement of context and mimicry in facial expression recognition.

The first adjustment involved revising the presentation order of the stimuli so that the context scene was presented prior to the expression. This method provided better control over an intended priming effect of context. In the previous experiments, simultaneous presentation of context and expression raised questions about the bi-directionality of an effect of context on perception of the expression. It was not certain that what was considered an effect of context was not due to an experimental artifact. With simultaneous presentation of the scene and expression, it was not certain that participants were not concentrating on the facial expression and not attending to the context scene. Sequential presentation would control for that possibility.

In addition to changing the presentation order, only two of the models were presented in this next experiment and the two selected displayed the widest range of measured expressivity. Finally, adjusting the design so that disrupted mimicry was a between group variable prevented any possible carry over effects that could diminish the effect of the manipulation. With these changes, the effect of context on, and a role for mimicry in, expression recognition was better understood.

6.2.iii Materials

For the current study, the basic paradigm as used in the two previously reported experiments was applied. The 24 pictures from the IAPS (Lang et al, 1999) identified in Chapter 4 (See Table 4.1) for the first experiment were used in the current experiment as context scenes. Corresponding facial responses to those 24 images by only 2 of the 4 models (1 female; F1 and 1 male; M2) were included in the paradigm.

Pilot information

Stimuli of different durations were tested on a pilot group of 30 participants. For this pilot, responses by 4 models to 8 context scenes were presented in both dynamic and static formats, each in 4 exposure times; 2000 ms, 1000 ms, 500 ms, and 250 ms. Eight scenes were chosen so that each scene would be paired as part of a same- and different-valence context pair. Because of the complications of multiple levels of exposure time but limited number context scenes there was concern about overexposure of model responses to too few scenes. It was decided that the paradigm should include the original 24 context scenes that had been presented in the first experiment (reported in Chapter 4) organized into same- or different-valence context pairs.

Different levels of recognition accuracy to static and dynamic expressions of 4 exposure times were confirmed during piloting. Levels of accuracy were highest on static and dynamic expressions of 1000 ms and lowest on dynamic expressions of 250 ms. The rate of accuracy to the 250 ms dynamic expressions was at 68%, which was statistically greater than chance ($p < .0001$, *Mean Difference* = .1800, 95% *CI* = .1542-.2058).

As a result of piloting, it was deemed unnecessary to present an expression for longer than 1000ms in order to perceive motion, and agreed that a difference in motion would be perceived in expressions of 1000ms and 250ms. Furthermore, a 250 ms expression would be perceived consciously and still contain evidence of motion. It was not the goal of the study to measure pre-conscious processing of facial stimuli. Finally, it was also considered necessary that participants see responses by each model in both exposure times. Consequently, the number of exposures and models was reduced to two instead of four, which allowed for a manageable number of trials within 1 hour of participation.

Face stimuli

Facial responses to the 24 context images were presented in static or dynamic format. In order to test for an effect of motion, each of the dynamic and static responses were presented in real time for 1000 ms and 250 ms. The expressions were not manipulated to increase or decrease speed of the expression, and therefore frames were not edited out of the 1000 or 250 ms length of the expression.[13] These responses were generated from the 5 s movies and static images used in the previous studies. On iMovie, each response was extracted from the original 5 s response. The apex of the expression was at the centre of the dynamic response, that lasted either 1000 ms or 250 ms. As before, the static image was extracted from the apex and presented for 1000 ms and 250 ms. Due to programming limitations for this experiment both dynamic and static expressions were presented in *dv* format.

Of the original four Model's, the two selected as stimuli for the current experiment had the most significant range in expressivity: Evaluations indicated a greater range of expressivity for Models F1 than F2 and for M2 than M1. Responses to 24 scenes by two different Models (F1 & M2) were presented in dynamic or static formats, in both 1000 ms and 250 ms exposure times. Thus, 192 expressions were the facial stimuli.

Context stimuli

Context scenes were organized into pairs consistent with the first experiment. Each facial expression would be presented with the actual scene to which it was the "true response" and with one other scene, the foil scene, to which it was the "false response".

---

[13] For comparable studies in which an effect of motion was measured, researchers presented facial expressions that had been edited, so that expressions were built from a compilation of frames that were not sequential in real time. These manipulations were intended to change the velocity, or speed of the unfolding of the expression.

The actual and foil scenes would make one "context pair". Twenty-four context scenes were sorted into 12 pairs. As reported for the first experiment, six pairs of context scenes were intended to elicit the same or similar emotion ("same valence" pairs) and six pairs were intended to elicit emotions that were opposite in valence ("different valence" pairs).

Organization of Stimuli

Stimuli were organized into two blocks with a repeat design for one of two possible sets of stimuli (Set 1, Set 2). In each block there were 96 pairings of a context scene and expression. Participants would see either Set 1 or Set 2 twice. Presentation order of stimuli was randomized within each block. This two set design was intended to prevent a possible transfer effect that appeared in the pilot data. Pilot data indicated that with a two- block design, exposure to the 1000 ms 'true' facial expression in the first block possibly affected (improved) performance on the 250 ms trials of same 'true' facial expression in the 2nd block of testing. So, in the final experiment, participants saw 24 expressions by two models in both the 'true' and 'false' response conditions, but *not* in both exposure times. For each pair of context scenes, subjects saw the true and false response of one of the two models in either 250 or 1000 ms, but not both in the same set (See Table 6.2). All context-expression pairs for each duration, involving the two models, occurred equally in the experiment through equal use of sets 1 and 2 across participants. However, each participant saw only one of these two sets twice in two blocks, in a different randomized order. So positive transfer in performance involving a particular context-model response pairing across blocks only occurred within a particular temporal duration and not across temporal durations

Table 6.1 Organization and controlled presentation of stimuli in two sets

Paired Context Scenes	**Model**	**True/ False**	**Duration(ms)**
Sample from Set 1			
2340 image	M2_2340	TRUE	250
9520 image	M2_2340	FALSE	250
9520 image	M2_9520	TRUE	1000
2340 image	M2_9520	FALSE	1000
9520 image	F1_9520_	TRUE	250
2340 image	F1_9520	FALSE	250
2340 image	F1_2340	TRUE	1000
9520 image	F1_2340	FALSE	1000
Sample from Set 2			
2340 image	M2_2340	TRUE	1000
9520 image	M2_2340	FALSE	1000
9520 image	M2_9520	TRUE	250
2340 image	M2_9520	FALSE	250
9520 image	F1_9520	TRUE	1000
2340 image	F1_9520	FALSE	1000
2340 image	F1_2340	TRUE	250
9520 image	F1_2340	FALSE	250

Table 6.1 illustrates with an example of one pair of context scenes how the true and false responses of a single pair were presented in different Sets. As evident in Table 6.1, the same participant did not see the true and false response conditions for each model in *both* exposure times. When one model's true response to a particular scene was presented in 0250 ms, the false response to that scene, involving that Model's, was presented in 1000ms. For the same participant, the true response to that scene was not presented in 1000ms and the false response was not presented in 250 ms. Another participant saw the reverse for the second model's response to the same context scene.

6.2.iv Experimental Design

In the current study the dynamic/static quality of facial stimuli and disruption of mimicry were between-group factors. Participants perceived either dynamic *or* static facial expressions and either experienced a pen manipulation intended to disrupt mimicry throughout the entire experiment, or they were free to mimic. Thus, the experiment adhered to a mixed factor design defined by 2 (Group; Static, Dynamic) x 2 (Mimicry: Disrupted or Free) x 2 (Video Time: 250 ms, 1000 ms) x 2 (Model's Response; true, false) x 2 (Congruence; context pairs that were congruent or incongruent in valence).

6.2.v Procedure

Presentation of scenes and expressions

Presentation of stimuli corresponded to Experiments 1 and 2 with the following changes; on every trial, a context scene preceded of a facial expression. Figure 6.1 shows the timing of a single trial, with the scene and expression each positioned in the centre of the iMacG5 computer screen. Simultaneous, side-by-side presentation of the scene and expression in the previous experiments had prohibited certainty that a participant's response to the context scene actually preceded their processing of the facial expression. By presenting the context scene first, it was certain that the participants' responses to context preceded their perception of the facial expression. Additionally, it meant that the longer exposure times of 5000 ms were no longer necessary. The duration of 5000 ms was initially intended to provide adequate time for participants to attend to the simultaneously presented context scene and expression.

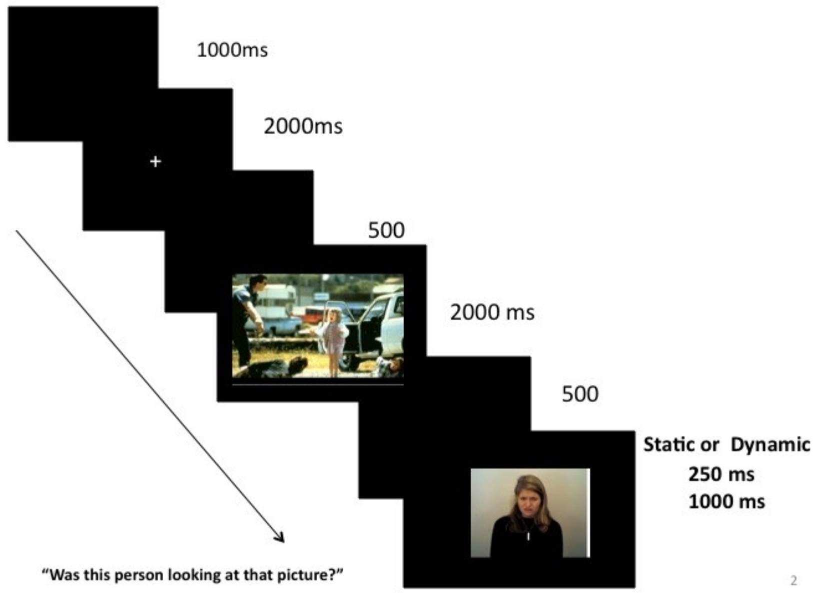

Figure 6.1 Final presentation times for a single trial ending with a response recorded after onset of facial expression.

The procedure followed the protocol for Experiments 1 and 2 with only minor modifications outlined herein. In the current experiment, the "bite-pen" manipulation that was intended to disrupt spontaneous mimicry was applied as a between-group factor. As before, participants saw either dynamic or static facial expressions.

During every trial, participants were instructed to respond "yes" or "no" to the question "Was the person looking at that image?" Instructions for which response keys to use were alternated. Half the subjects entered "z" for "yes" and "/" for "no", while half the subjects pressed "/" for "yes" and "z" for "no". Response key instructions were coded as well as subjects' handedness. There was equal number of left-handed responders in both conditions of response keys. Alternating response keys was intended to provide better control over the possibility of participants' motor hand dominance influencing results by responding impulsively with the dominant hand.

The third and final change to the procedure was that participants could response immediately following stimulus onset. As reported above, because the context scene and facial expression were no longer presented simultaneously, it was no longer necessary to provide ‘adequate time’ for processing both images and recording participant response after stimulus offset. Participants were instructed to respond as quickly as possible without forfeiting accuracy. It was anticipated that responses were not affected by waiting for the stimulus offset.

Participants were quasi-randomly assigned to 1 of 8 counterbalanced orders defined by Set 1 or 2, Static or Dynamic faces, pen or no pen manipulation, and response keys. Following detailed instruction and obtaining participant’s consent, the experimenter sat with the participant during 10 practice trials, which were not recorded, and left the room during actual data collection. Practice trials were followed by 2 blocks of 96 trials for which each response was recorded.

6.2.vi Data analysis

Primary variables of interest

The dependent variable of interest was accuracy. Correct responses were determined by correctly identifying the Model’s facial expression as the true or false response. A response to each trial was coded as either correct = 1 or incorrect = 0.

Data were sorted into groups defined by dynamic or static facial expressions, and then by disrupted mimicry or free to mimic. Of the 32 participants exposed to dynamic expressions, 16 experienced the pen manipulation intended to disrupt mimicry and 16 were free to mimic. Of the 31 participants who were exposed to static expressions, 16 experienced the manipulation and 15 were free to mimic.

As with the first experiment, context scenes were organized into pairs that were either were congruent in valence (both positive or negative) or incongruent in valence (positive and negative). The variable was labeled Congruence of Context Pairs. The variable Model's Response indicated when the Model's expression was the true or false response to the context image. The variable Video Time had two levels for the duration of a facial expression as either 250 ms or 1000 ms. Static and dynamic expressions were presented in both durations.

The mean proportion of correct responses were calculated for every subject and submitted to a mixed analyses of variance (ANOVA), for which the main 2x2x2x2x2 model consisted of "Congruence of Context Pairs", "Model's Response", "Video Time" as within-subject factors and "Group" and "Disrupted Mimicry" as between-group factors. Preliminary analysis of variance was conducted on control factors of "order" and "response keys".

Subsequent analyses were run on data by the static and dynamic groups. *A priori* expectations for differences in accuracy to static and dynamic expressions when durations were 1000ms and 250 ms, and when mimicry was disrupted prompted the select analyses.

Analyses of variance were conducted using R version 2.15, with application of the "*ez* package", version 4.2.2, (Lawrence, 2009).

Effect size and Outliers

As with the previous experiments, the size of an effect was reported as general *eta* squared, which would reliably measure the proportion of total variance in a dependent variable explained by the independent variables. Alpha levels were set at a criterion of .05 for significance. Unless otherwise stated, paired *t* tests or one-way ANOVAs were conducted to test for simple effects. Errors bars represent the FLSD unless otherwise stated.

For Accuracy a criterion for subject outliers was set at any subject mean that was three standard deviations beyond the group mean. Subject means beyond that criteria would be considered an outlier and their data were dropped. The data of two participants were dropped because of extreme biased responding that affect accuracy; One participant judged every response as False and one other judged most responses as True.

Thus, the data of 63 participants were analyzed. Seventy-six percent of participants were female and the mean age remained at 20 years, with a range of 18-39 years. Five left-handed participants were evenly distributed in the groups defined by exposure to static or dynamic expressions and whether or not they were free to mimic.

6.3 Results

A preliminary analysis of variance on control factors "Order" for order of set of stimuli, and "response keys" showed no significant main effects or interactions with either factor.

The ANOVA table for the main model for Accuracy is included in the Appendix. Due to the multiple comparisons, the risk of a Type I Error was considerable. Therefore, the criterion for significance was adjusted by Bonferroni corrections. The criterion of .05 was divided by the number of significance tests (31) generating a more strict criterion of p = .0016.

Collapsed across between-group factors of Group (Static/Dynamic expressions) and Disrupted Mimicry (Pen or NoPen manipulation), the total mean rate of accuracy was 70% (SD = .04) (N = 63) with a range of 60-82%. All participant means were above chance. This rate of accuracy is consistent with results in Experiment 1, (70%/05) and is significantly different from chance, p < .0001, *M Difference* = .2000, 95% *CI* = .1840-.2160. Experiments 1 and 3 were comparable in their use of both congruent and incongruent-valence pairings of context scenes.

6.3.i Was there a dynamic advantage?

Results for Experiment 3 supported the hypothesis for a dynamic advantage. Participants were significantly more accurate in their judgments about dynamic compared to static expressions, $M_{Dynamic}$ = .73, SD = .04, M_{Static} = .68, SD = .04, F(1,59) = 20.615, p <.00003, ${}_{gen}\eta^2$ = .02.

In addition to anticipated higher levels of accuracy to dynamic than static expressions, there was an expectation of higher levels of accuracy to 1000 ms dynamic relative to 250 ms dynamic expressions. The same difference was not expected for static expressions. The means in Table 6.2 show significantly higher levels of accuracy when dynamic expressions were 1000 ms than 250 ms, $F_{Dynamic}$(1,30) = 37.959, p <.00001, ${}_{gen}\eta^2$ = .09. In comparison, levels of accuracy to static expressions of 1000 ms and 250 ms did not differ (p = .80). Higher levels of accuracy to dynamic expressions of 1000 ms suggest there was an advantage in recognition accuracy when dynamic expression long rather than brief. The means indicate participants were no more accurate in their judgments about dynamic expressions of 250 ms than static expressions of 250 and 1000 ms, however, with a corrected criterion for significance of .0008, the interaction of Group with Video Time was not significant, F(1,59) = 8.464, p = .005, ${}_{gen}\eta^2$ = .016. Better accuracy when dynamic expressions were long supported the hypotheses of additional information available in the longer dynamic expression, such as motion, explaining better recognition when expressions are dynamic than static.

Table 6.2. Levels of accuracy as a function of Group and VideoTime (N=63)

	Accuracy % (M/SD)	
	Dynamic	Static
1000 ms	.76/.05*	.68/.07
250 ms	.69/.05*	.69/.07

*Significant at .003

6.3.ii Was there an effect of context on processing dynamic and static facial expressions?

Significantly higher levels of accuracy to judge the model's response as true than false when the context scene was congruent in emotion valence was identified *a priori* as an index of an influential effect of context. Additionally, an influence effect of context was expected to be greater when expressions were dynamic than static. Consistent with Experiments 1 and 2, the hypothesis of an influential effect of context on recognition was supported.

A main effect of Congruence showed that participants were always more accurate when scenes were of an incongruent pair, $F(1,59) = 230.055$, $p < .00001$, ${}_{gen}\eta^2 = .19$. Consistent with Experiment 1, the difference in accuracy was explained by easy recognition of the false response to an emotionally incongruent foil scene, which was at ceiling levels (88%). In all other conditions, when the expression and scene were congruent in valence, it was difficult to recognize the model response as true or false and levels of accuracy were lower. Figure 6.2 shows the significant interaction of Model Response and Congruence of Context Pairs, $F(1,59) = 354.038$, $p < .00001$, ${}_{gen}\eta^2 = .19$. Simple effects confirmed that when scenes were incongruent in emotion valence, participants were more accurate recognizing the false than true response, $F(1,62) = 108.505$, $p < .00001$, ${}_{gen}\eta^2 = .52$. In comparison, when scenes were congruent in emotion valence, participants were significantly more accurate when they judged the model response as true than false, $F(1,62) = 6.919$, $p < .01$, ${}_{gen}\eta^2 = .08$.

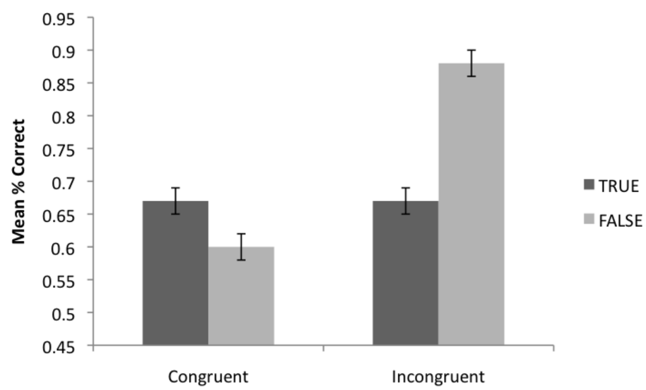

Figure 6.2 Model Response and Congruence of Context Pairs interacted, $p < 0001$, to affect better recognition accuracy of the false than true response with incongruent scenes, $p <.0001$, and the true than false response with congruent scenes, $p = .01$. $N = 63$, Error bars are FLSD.

Consistent with results in Experiment 2, the influence of the congruent scene on judgments of the expression as true was stronger when expressions were dynamic compared to static. The interaction of congruence of context scenes and model response accounted for 32% total variance in accuracy to dynamic expressions, $F(1,30) = 245.017$, $p < .00001$, ${}_{gen}\eta^2 = .32$, and 16% of total variance in accuracy to static expressions, $F(1,59) = 122.781$, $p < .00001$, ${}_{gen}\eta^2 = .16$. When the context scenes were congruent, participants were significantly more accurate judging dynamic expressions as true than false, $F(1,31) = 21.718$, $p = .00006$, ${}_{gen}\eta^2 = .36$. In comparison, when the context scene was congruent there was no difference in recognition of the static expression as true or false, $p = .83$, ${}_{gen}\eta^2 = .00$. Higher levels of accuracy to recognize the true than false response when the scene was congruent in valence supported the hypothesis of an influence of context on expression recognition. Moreover, the influence affected recognition of dynamic but not static expressions.

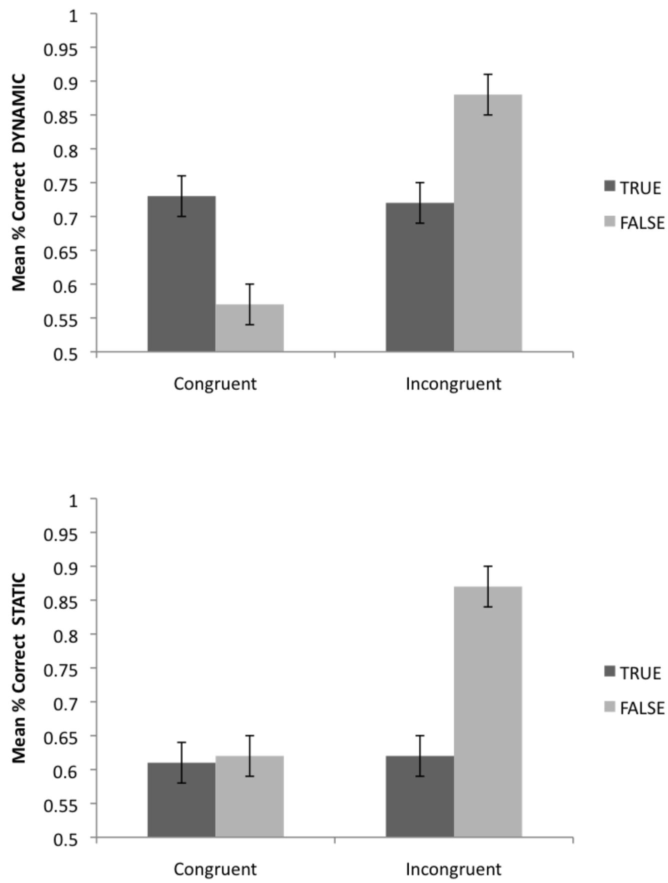

Figure 6.3 Model's Response interacted with Congruence of Context Pairs when expressions were Dynamic (upper) *p* < .0001, and Static (lower) *p* < .0001. When scenes were of incongruent pairs, expressions were better recognized as false than true. When scenes were congruent, Dynamic expressions were more accurately judged as true than false, *p* = .00006. n_{Dynamic} = 32, n_{Static} = 31, Errors Bars are FLSD.

The influence of context did not differ with exposure time. Video Time did not significantly interact with Model Response, Congruence, or static vs dynamic expressions. Thus, the hypothesis for a greater influence of context on dynamic 1000 ms than dynamic 250 ms expressions was not supported.

6.3.iii Does mimicry underlie recognition of dynamic and static expressions?

Consistent with Experiments 1 and 2, there was no main effect of Disrupted Mimicry on recognition, p = .23. Disrupted Mimicry did not directly affect accuracy to either static or dynamic expressions, p = .18. In the current experiment, Disrupted Mimicry interacted with VideoTime and Congruence of Context Scenes, $F(1,59) = 12.319$, p = .0008, $_{gen}\eta^2$ = .01. Figure 6.4 illustrates the interaction.

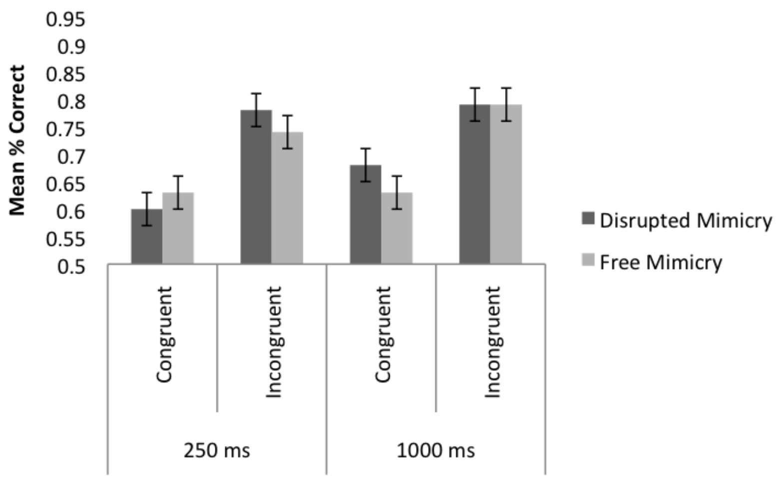

Figure 6.4 The interaction of Disrupted Mimicry with Video Time and Congruence of Context pairs, p = .0008. When context scenes were congruent disrupted mimicry facilitated recognition of 1000 ms expressions, p = .013, $_{gen}\eta^2$ = .10, but not 250 ms expressions, p = .15, $_{gen}\eta^2$ = .03. Simple effects were not significant on 250 ms expressions. N = 63, Error bars are FLSD.

As shown in Figure 6.4, the effects of disrupted mimicry were weak and differed with duration of expressions. Simple effects confirmed a significant facilitation effect of disrupted mimicry on recognition of 1000 ms when context scenes were congruent, $F(1,61) = 6.556$, p = .013, $_{gen}\eta^2$ = .10. In comparison, disrupted mimicry had weak effects on recognition of brief, 250 expressions that differed with congruent and incongruent context scenes and differed from 1000 ms expressions. When context scenes were congruent in valence, the means indicated lower levels of accuracy to 250 ms expressions when mimicry was disrupted than when mimicry was possible. Whereas, when context scenes were incongruent, the means indicated higher levels of accuracy with disrupted mimicry. However, these differences in accuracy to 250 ms expressions did not reach significance.

There was an *a priori* expectation for impaired recognition of static expressions when mimicry was disrupted, and possible impairment effect on recognition of dynamic expressions of 250 but not 1000 ms. Figure 6.5 illustrates the lack of effect of disrupted mimicry on static and dynamic expressions of 250 and 1000 ms.

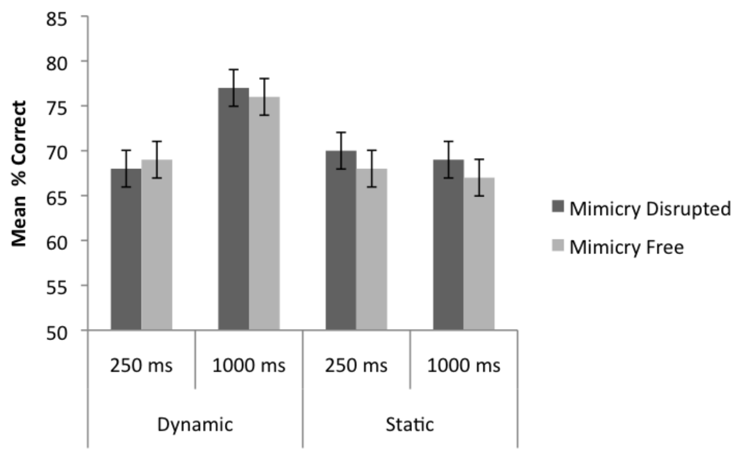

Figure 6.5 Recognition accuracy is to dynamic and static expressions of 250 and 1000 ms is not significantly impaired or facilitated by disrupted mimicry. Error bars are FLSD.

Figure 6.5 shows better recognition when dynamic expressions are 1000 ms than 250 ms or than static expressions of 250 and 1000 ms, which is not affected by disrupted mimicry. Figure 6.5 illustrates there was no significant effect of disrupted mimicry on accuracy to dynamic or static expressions that were 250 and 1000 ms. The means indicate slightly higher levels of accuracy to static expressions when mimicry was disrupted, but the effects are not significant. Thus the hypothesis for impaired recognition of static expressions and dynamic expressions of 250 ms caused by disrupted mimicry was not supported.

6.4 Discussion

In the current experiment, participants were more accurate to process judgments about dynamic than static expressions. Increased effects of context and mimicry relative to the previous experiments were anticipated due to a modified experimental design and procedure. Modifications were motivated by the goal of obtaining greater control over the variables of interest and thereby improving sensitivity in measured responses. For example, in the current design, a context scene was presented prior to a facial expression to erase any possibility that participants might have attended to the model's expressions prior to the simultaneously presented context scene. Presenting the scene first eliminated that possibility and provided a stronger measure of the effect of participants' response to the context scene on their perception of a following facial expression. Additionally, Disrupted Mimicry as a between-group variable was intended to avoid the possibility of any carryover effects. In Experiments 1 and 2, it was uncertain if a lingering effect of the pen manipulation carried over to alternating blocks of trials in which participants did not experience the manipulation. With these modifications, stronger effects of context and disrupted mimicry on processing different qualities of facial expression were expected, relative to results observed in the previous experiments. In this behavioral experiment, quality of facial expression was operationalized as static or dynamic expressions as well as brief and long duration times. Processing had been measured as accuracy to correctly recognize expressions as the true or false response to a preceding contextual scene.

In order to obtain greater insight into differences in processing perceived static and dynamic facial expressions, exposure times were modified so that each expression was presented for 1000 ms and 250 ms. The shorter duration of dynamic expressions was expected to negatively impact judgments, because brief expressions of 250 ms contained less dynamic information, such as less perceived change in the expression. Perception of change was identified as an important aspect of a dynamic expression that is not available in a static expression. Moreover, motion was identified as inherent in the perception of change (Ambadar, 2005). Motion conveyed by changing temporal-spatial relations in facial features contributes to the complexity of a dynamic expression, and is expected to assist in accurately recognizing the meaning of a given expression (Thournton & Kourtzi, 2002). The question addressed in this study was whether the brief dynamic expression would contain sufficient motion cues to support more accurate recognition than that of a static expression. The results showed that recognition to brief dynamic expressions was not more accurate than to static expressions.

Participants were significantly more accurate in their judgments about dynamic than static expressions but only with dynamic expressions of longer duration. Motion cues depicting change in spatial temporal relations of the facial features have been shown in several studies to trigger activation of systems that support dynamic but not static processing (e.g., Kilts, et. al., 2003; Trautmann, 2009). Early activation of the aSTS and pSTS, associated with perception of change in the eyes and mouth and biological motion (Pitcher et al., 2011). Nonetheless, the amount of information available in the brief dynamic expression presented in the current work was not enough to support more accurate recognition of dynamic relative to static brief expressions. Perception of motion cues in the few frames of a 250 ms expression were not enough to depict the unfolding of the expression that would support accuracy to the task in the current work. Hence, the dynamic advantage was not was not found in brief dynamic expressions.

All expressions were centred at the apex of the expression and the onset and offset of dynamic expressions evolved toward and away from the apex. With 125 ms on either side of the apex, brief dynamic expressions were similar to the static expressions. Although some theorists proposed that no additional information was needed beyond the cues available in the apex of a static expression (Bould et al., 2008), the results here indicate otherwise. Better accuracy to long dynamic expressions suggests there is additional information in the offset and onset of the expression to support recognition. Naturally unfolding dynamic expressions of 1000 ms, contained twice as much information as the brief expression, with 500 ms on either side of the apex. Long expressions revealed more of the unfolding or change in an expression. As a result, participants were more accurate with 1000 ms dynamic expressions than 250 ms dynamic expressions.

The two duration times were not expected to affect judgments about static expressions. No overall difference in recognition of 1000 ms and 250 ms static expressions was consistent with that expectation. The results indicated that static expressions of longer durations were processed no differently than brief static expressions.

The results supported the hypothesis of a different quality of information available in long dynamic expressions that is not available in static expressions, that contributes to accurate recognition of expressions. Perception of change is one example of such information that explains better accuracy in recognition of dynamic than static facial expressions. In the current work, dynamic information supported participant's ability to discriminate natural facial expressions as the appropriate response, or not, to a given context scene.

Based on the findings from Experiments 1 and 2, a greater effect of context on perception of dynamic more than static expressions was expected in the current experiment. An effect of context was measured as significantly higher levels of accuracy for judgments of the Model's expression as the true than false response when context scene and expressions were congruent in emotion valence. Judgments of the response as true implied the response matched the participant's response to the given context scene. When scenes and expressions were congruent, participants were indeed more accurate when judging dynamic expressions as the true rather than false response. This effect was not observed in response to static expressions.

Unless static expressions were presented with a context scene of a different valence emotion, participants did not discriminate well between true and false responses. Unlike judgments about dynamic expressions, participants who saw static expressions were not likely to judge the uncertain response as true. Only in the easy condition when the valence of scenes and expressions were incongruent, did participants accurately judge the static expression as false. Hence, there was no immediate evidence of an effect of context on judgments about static expressions.

The results in Experiment 3 did not show significantly different effects of disrupted mimicry on recognition to static and dynamic expressions. Disrupted mimicry facilitated recognition of long expressions when context scenes and expressions were congruent in valence. The effect was general and not specific to either static or dynamic expressions. The results did not indicate a significant effect of disrupted mimicry on recognition of expressions the true or false response.

Due to anticipated reduced motion cues in a brief relative to long dynamic expression, it had been suggested that brief dynamic expressions would benefit by spontaneous mimicry. Similar to static expressions, additional motor activity generated by facial mimicry might benefit the simulation process of brief dynamic as well as static expressions. The weak and not significant effect of disrupted mimicry on 250 ms expressions was not specific to static or dynamic expressions. Since there was also a general, but significant, facilitative effect on 1000 ms expressions that was not specific to dynamic or static expressions, perhaps the effects of mimicry or disrupted mimicry are better understood as a function of exposure time rather than the static or dynamic quality of facial expressions.

In the current study, a dynamic advantage in accurate recognition was observed to dynamic expressions of 1000ms, but not 250 ms. That the advantage was lost on brief dynamic expressions supported the hypothesis for perceived motion as a critical element in processing dynamic expressions. Context continued to have a strong influential effect on dynamic but not static expressions. An weak facilitative effect of disrupted mimicry was observed on 1000 ms expressions that was not specific to static or dynamic expressions.

CHAPTER 7: GENERAL DISCUSSION

The objective of this thesis was to use a behavioral measure to investigate the underlying processes for facial expression recognition. The overarching goal was to test simulation theory as a valid explanation for situated facial expression recognition. To that end, the research was built on the premise that perception of another person's facial expression leads to the observer experiencing a corresponding, simulated emotional state that is concurrent with activation of multiple systems (e.g. sensorimotor, affective, cognitive) involved in experiencing the emotional state. The shared emotional state is implicitly attributed to that other person (Gallese & Goldman 1998; Gallese, 2003, 2007; Goldman 2005), and thereby supports recognition.

To increase ecological validity in a testing paradigm for expression recognition it was deemed important to include stimuli of natural dynamic expressions, as opposed to artificially generated expressions. In addition to increasing ecological validity, the inclusion of context was considered a critical factor for testing simulation theory, as it added an additional level of simulation that was seldom accounted for in the recognition literature but is part of the recognition process as it occurs in day-to-day social interactions. Based on theories of embodied cognition, an observer's understanding of perceived events is informed by their concurrent multi-modal response to contextual factors. For example, during a social interaction, recognition of another person's facial expression is informed not only by the signal emitted from the other person's facial activity, but also the observer's multimodal response to the surrounding social context. Based on increasing evidence for an influential effect of context on perception of facial expressions (Aviezer, et al., 2008; Barrett, Mesquita & Gendron, 2011; Halberstadt, et al., 2009; Righart & de Gelder, 2006, 2008) one of the main hypotheses in this thesis was that an observer's embodied response to a shared context would influence his or her judgments about the observed facial expression. To account for only the facial expression as the source of information on which the observer bases their judgment about the emotional state of their social partner is omitting another important source of information that has a significant contribution to the judgment. Specific to the task in this thesis, it was hypothesized that participants' embodied response to a presented context scene would influence their judgment of the model's response as appropriate or not to the given scene that was of the same emotion valence.

Another main hypothesis for this research was of facial mimicry as an important underlying mechanism in expression recognition. Like the embodied hypothesis for context this hypothesis was consistent with simulation theory, and was influenced by seminal research in which disruption of mimicry was shown to impair expression recognition (Niedenthal et al., 2001; Oberman et al. 2007). Based on theories that propose shared representation for perception and action, coding for perception of the expression overlaps with coding for generating an expression (Prinz, 2005). In addition to low levels of activity in the motor regions in the brain due to perception, mimicry is a source of generated motor activity. The generated motor activity contributes to a stronger motor signal, which enhances downstream simulation to assist in recognition of the facial expression (Goldman & Sripada, 2005; Iacoboni, 2008). What was new in this research was testing mimicry as a causal factor in *situated* recognition of natural expressions. The mimicry hypothesis had been supported by blocking mimicry during perception of static and computer-generated expressions, but not tested on responses to natural expressions that were presented with scene context.

The first Experiment, reported in Chapter 4, addressed two hypotheses regarding context and mimicry. Except when the expression was an obvious false response to a foil context scene that was incongruent in valence, higher levels of accuracy for recognition of the response as the true than false were expected. The expected difference was explained by the predicted influence of context. When expressions and context scenes were the same or similar emotion, the participant's embodied response to the context scene influenced perception of the expression and increased the likelihood of judging the expression as matching to their own – as the true response to the context scene. Significantly higher levels of accuracy to judge the response as true on congruent valence trials was identified *a priori* as a measure of an influence of context. The results confirmed a strong effect of context. Furthermore, effects of context were reliably found throughout the three experiments. How context interacted with other variables to influence judgment is explained below.

The second objective in Experiment 1 was to test for a role of mimicry in recognition of natural dynamic expressions by disrupting mimicry. Two manipulations intended to disrupt mimicry in previous seminal research (Niedenthal et al. 2001; Oberman et al. 2007) were applied in Experiment 1. The results in the first experiment did not support that hypothesis for impaired recognition due to a disruption of mimicry.

A third main hypothesis for this research arose from results in the Experiment 1. Dynamic expressions were presented as testing stimuli because of the motivation for increased ecological validity. However, results of interfering with mimicry reported in Experiment 1 differed from reports in the literature that were based on testing with static or computer generated expressions. To determine if the dynamic quality of stimulus explained the difference in results, testing in Experiments 2 and 3 was conducted on dynamic and static facial expressions. Reviews of the literature on the dynamic quality of facial expressions lead to the hypothesis of a dynamic advantage – more accurate expression recognition of dynamic than static expressions. Although a question of different underlying processes for static and dynamic expressions was not an initial objective for the research, it presented itself from the research and influenced the work in Experiments 2 and 3. Neuroimaging studies had reported differences in neural activity associated with processing dynamic and static expressions (Arsalidou et al., 2011; Kilts et al., 2003; LaBar et al., 2003; Pitcher et al., 2011; Sato et al., 2004; Trautmann et al., 2009; 2013), and several reports cited differences in covert mimicry and intensity ratings to perceived dynamic and static expressions (Achiabou et al, 2008; Ryzmaryk et al., 2011; Sato & Yoshikawa, 2004, 2008). However, there was no information for directly comparing the simulation processes for each quality of expression. I was not aware of any comparative study that had tested for a causal effect of mimicry on recognition of natural dynamic and static expressions that, moreover, were presented with scene context.

Presenting both dynamic and static expressions provided an opportunity to learn more about the influence of context and the role of mimicry in recognition of each. Several studies presented evidence for higher levels of motor activity associated with processing static relative to dynamic expressions, which lead to their hypotheses of a predominant reliance on the motor system for recognition of static expressions (Kilts et al., 203; Trautmann et al., 2009; 2011, see also Arsalidou, Morris & Taylor, 2011 for a review). The same studies also identified more widespread activation associated with processing dynamic expressions, which lead to the hypothesis in this research for less reliance on motor mimicry for recognition of dynamic relative to static expressions. The mirror neuron system had been associated with expression recognition based primarily on research with static expressions (e.g. Carr et al., 2003). Because areas of neural activity associated with processing dynamic expressions were not limited to predominantly the visual and motors areas, as they were for static expressions (Arsalidou, Morris & Taylor, 2011; Kilts et al., 2003; Trautmann, et al., 2009, 2011), it was suggested in this research that processing dynamic expressions was less reliant on the motor system compared to processing static expressions. This perspective led to a hypothesis of reduced need for mimicry in recognition of dynamic compared to static expressions. The second experiment continued to test for an effect of context and disrupted mimicry on recognition. Recognition was tested to dynamic and static versions of the same facial expressions.

Chapter 6 referenced evidence in the literature for mimicry to dynamic expressions that was based on expressions of long durations whereas much of the evidence for mimicry to static expressions was based on long and short durations (mostly short). The lack of information about mimicry to brief durations of dynamic stimuli raised a question in this research of how processing brief dynamic expressions would compare to processing static expressions. Perceived motion cues had been identified as a key component in a dynamic expression to explain a dynamic advantage in recognition. Motion cues enable the perception of change in a dynamic expression that is not available in static expressions (Ambadar et al., 2005). Experiment 3, reported in Chapter 6, tested a hypothesis of perceived motion as a causal factor in the dynamic advantage in expression recognition. It was predicted that a significant reduction of motion cues would disable the dynamic advantage in expression recognition and effect little difference in accuracy to brief dynamic expressions relative to static expressions of brief and long durations. Furthermore, it was hypothesized in this thesis that less motion perceived in brief dynamic expressions increases the demand to generate motion by mimicry. Thus, a difference in recognition was expected when dynamic expressions were long and brief expressions and greater similarity in recognition was expected between brief dynamic and static expressions. In Experiment 3, dynamic and static expressions were presented in durations of 1000 ms and 250 ms. The main objective was to compare recognition of 1000 and 250 ms dynamic expressions and 250 ms dynamic with static expressions. Additionally, the results from Experiment 1 and 2, based on recognition of 5000 ms expressions, were compared with the findings for the 250 ms and 1000 ms durations.

7.1.i The influential effect of context on recognition

An influential effect of context was consistently observed on judgments about dynamic expressions. In the three experiments, recognition of the model's response as true was far superior to recognition of the response as false when foil context scenes were congruent in valence. The left panel of figure 7.1 below summarizes the influence of context on dynamic expressions observed in Experiments 1, 2, and 3. The same effect was observed on all three durations of dynamic expressions.

A difference in the magnitude of the effect of context between the 1000 and 5000 ms dynamic expressions can be partially explained by changes to procedure and stimuli in Experiment 3. Presenting the context scene and expression sequentially in Experiment 3 resulted in a strong priming effect of the context scene on judgments of the facial expression. A smaller effect was evident on simultaneous presentations in Experiment 1 and 2 that strengthened with a sequential presentation in Experiment 3, at least on 1000 ms expressions.

Changes to procedure helped to account for the increased influential effect of context on 1000 relative to 5000 ms dynamic expressions. It did not explain the difference between 1000 and 250 ms durations in Experiment 3 in which both durations were presented following the same sequential procedure. Recognition of the 250 ms model response as true was superior to recognition of the false, however, the effect was weaker than observed on 1000 ms duration. The result indicated a difference in processing between dynamic expressions of long and brief durations.

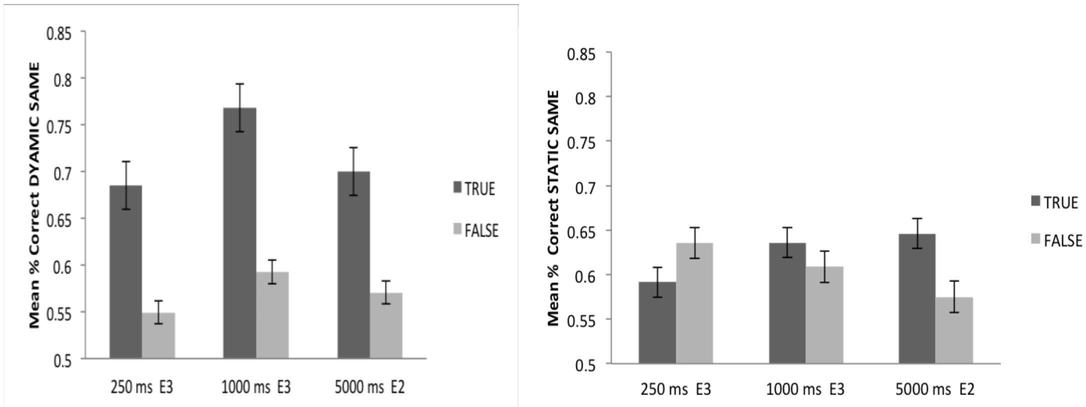

Figure 7.1 Accuracy to true and false responses on same valence trials as a function of static and dynamic expressions and duration of expression, across three experiments.

The influential effect of context on 250 ms dynamic expressions also indicated a difference in processing brief dynamic and static expressions. In response to 250 ms static expressions, there was no effect of context on recognition (Figure 7.1, right panel). As reported in the text, recognition of the false response was high when the static expression was presented with a context scene that was incongruent in valence. Otherwise, on all other trials with 250 ms static expressions and congruent context scenes there was no accurate discrimination of the true from false response. Recognition of the false response always remained higher than the true response. The absence of an effect of context on brief static expressions was one example of how the two qualities of brief expressions, static and dynamic, differed in regard to an influence of context.

In response to longer static expressions, when the context scene was congruent in valence, recognition of the true response was better than recognition of the false response and the effect was strongest when static expressions were 5000 ms. Duration was a factor that interacted with context to affect recognition to static expressions. Overall, an influence of context on expression recognition was greater on dynamic than static expressions and greater on longer than brief static expressions.

7.1.ii Evidence for a dynamic advantage

A differential effect of context may contribute to an explanation for better accuracy in long dynamic compared to brief dynamic, and static expressions. Better recognition accuracy was observed on dynamic expressions of 5000 ms and reached significance on 1000 ms compared to static expressions. Hence, the hypothesis for a dynamic advantage was supported, when expressions were long in duration. Moreover, this advantage is restricted entirely to an advantage on true expressions, as shown in Figure 7.1. The longer exposure times improved performance on true expressions in the dynamic compared to static condition, while not providing any relative gain for false expressions. Accuracy was less for dynamic 250 ms expressions, and overall no better than static expressions, which indicated a loss of valuable, dynamic information when the duration of a dynamic expression was reduced. The findings indicate that the dynamic information is quantifiable, in that less of it perceived in a brief expression reduced the dynamic advantage.

The greater influential effect of context when expressions were dynamic relative to static may be partially explained by the more expansive neural activity associated with processing dynamic relative to static expressions. It was proposed in the thesis that the simulation of dynamic expressions involves multiple systems that overlap with systems associated with processing the context scene. Furthermore, with the sequential presentation of the context scene preceding the expression, processing context was already active when processing the dynamic expression was initiated. An overlap with systems already active likely resulted in an integration of the represented context scene into the represented meaning of the perceived facial expression.

Reduced accuracy and strength of the effect of the context on judgments about 250 ms relative to 1000 ms dynamic expressions were indications of a reduction in dynamic processing of the 250 ms expressions. The expansive, multi-system processing of the 1000 ms dynamic expressions was diminished for 250 ms durations. The brief duration was intended as a measure of reduced motion that would typically enable perception of change in the dynamic expression. If indeed motion is the critical component in the dynamic advantage, then these results would suggest that early detection of motion in the facial expression triggers subsequent activity to support neural processing of dynamic expressions and without adequate perceived motion, less activity to support the dynamic processing is triggered.

Additionally, in the everyday ecological situation with dynamic expressions, there is a strong expectation that another person's expression will be congruent with one's own in a given, shared situation, partially due to conceptual structures such as schemas and scripts. With that expectation is a bias to interpret the other's expression as congruent with the situation. In the current task, this bias occurred for all dynamic durations. In dynamic expressions, perceivable cues indicating congruence are continuously revealed as the expression unfolds over time. The response to a static, unchanging expression is different. It is a non-natural stimulus with a peak expression of an emotion that may or may not correspond to the context scene. Without motion cues, nor the full range of emotional responsivity in a dynamic expression, this task may rely more on mimicry of the most salient expressive cues. The observer searches for cues in their motor activity and simulation of the expression. This suggested process is similar to one proposed by Goldman & Sripada (2005), although they did not specify the process as particularly *supplementary* of a static stimulus. The findings from the research for this thesis suggest that with longer durations, there is more time for the simulation process to continue, and cues of the expression that appear to be congruent with the context become clearer.

7.1.iii Effects of disrupted mimicry on expression recognition

There was an a priori expectation for a greater impact of disrupted mimicry on static than dynamic expressions. The effects of mimicry were not straightforward. Ultimately, the effects were weak and differed with duration more so than the dynamic or static quality of a facial expression.

The effect of disrupted mimicry was neither unidirectional nor robust. An unexpected facilitative effect of disrupted mimicry occurred on 1000 ms expressions. The effect did not significantly differ when expressions were dynamic and static. When expressions were 250 ms a trend for reduced accuracy when mimicry was disrupted was observed that was also not specific to either dynamic or static expressions. These effects on 1000 and 250 ms expressions occurred when context scenes were congruent in valence.

Ultimately, in this thesis, the hypothesis of mimicry underlying recognition of expressions was not supported when expressions were dynamic or static. Instead, there is a suggestion that mimicry may facilitate recognition of expressions with longer durations, and when expressions were difficult to discriminate as the true or false response to a context scene of the congruent valence.

Oberman and colleagues (2005) used static prototypical expressions in a forced choice paradigm. The current study presented more complex facial stimuli and asked a question demanding explicit judgments beyond categorization. The judgment required rapid detection of features in the facial expression as well as an attributional judgment about the expression as a possible response the given context scene. Hence, in addition to efficient feature detection, the task likely activated conceptual structures and inferential processes in addition to motor activity. As outlined in the model presented in Chapter 3, it is likely that performance on the task engaged processing of both the mirror neuron system and the mentalizing system. With the engagement of both systems that show greater overlap with processing dynamic expressions, the nature of the task itself may partially explain the superior performance on dynamic expressions. The need to form an explicit judgment, while experiencing the influence of context on processing a facial expression is likely associated with activation of affective, cognitive and rewards systems that are in addition to the motor, visual and somatosensory systems, especially during processing of dynamic expressions. Furthermore, the described processing would also entail the formation of inferences about one's own response to the scene as well as the perceived expression as a response to the scene. Formation of such inferences, that may or may not reach awareness, are associated with activation in the mentalizing system. How critical mimicry is to this complex series of processes was questioned. The results in this research suggest mimicry is not essential to the process.

The difference in strength and direction of the disruption effect compared to reports in the literature raised obvious questions about the reliability of results in the current. A weakness in the experimental design was the lack of a control condition. There was no inclusion of a cognitive or motor task that would act as a control, thus there remains some uncertainty if the effect of the pen manipulation is indeed due to a disruption in the motor component of the simulation process, or due to cognitive load.

Similarly, there was uncertainty if the effect of the manipulation was not due to attention, particularly the facilitation effects resulting from the manipulation that was intended to disrupt mimicry. In the facilitation effect, recognition was better in the disrupted mimicry condition. Responses in conditions of disrupted mimicry often showed less variance in accuracy compared to the free to mimic condition. However, if the effects of the pen manipulation were due to heightened attention then it seems unlikely that recognition would differ to true and false responses and vary as a function of dynamic and static expressions. Thus, attention was not an immediate explanation for the unexpected facilitation effect.

Even more beneficial would be the direct comparison of expression recognition to dynamic and static expressions without context scenes. The inclusion of a control condition of no context would undoubtedly help to clarify the role of mimicry in expression recognition of static and dynamic expression that were presented with and without context.

7.1.iv Implications

The effects of mimicry are not straightforward. Disrupted mimicry had a weak effect on processing both static and dynamic expressions, but the effect was not strong and not disruptive. The impact of mimicry on recognition of natural expressions presented with scene context may be better understood by factors of duration and context than by differences in dynamic and static qualities.

Consistent effects of scene context on recognition provided strong support for the hypothesis of an influential role of observer's response to the context scene in their processing of the facial expression. The results were a strong indication of context as important and influential component of the perceptual experience during a social interaction. Importantly, the effect of context was much greater on dynamic than static expressions. Clearly context is processed differently when presented with static and dynamic emotional expressions of varied duration.

The findings for context emphasize difference in processing dynamic and static expressions. A dynamic advantage in situated recognition was supported, in response to natural expressions of long duration. The reduced levels of accuracy to dynamic expressions of 250 ms suggest a loss of quality of information in the expression that impacted recognition. The brief duration was intended as a measure of reduced motion that would typically enable perception of change in the dynamic expression. The manipulation of motion in this study, although more natural than manipulations of velocity reported in the literature, did not allow for precise control over the amount of motion that could be perceived. Consequently, it is not certain that a lack of motion cues was the absolute explanation for reduced levels of accuracy in the 250 ms dynamic expression. What seems certain is that longer durations of emotional expressions greatly increases accurate performance on dynamic compared to static expressions, and that this improvement in performance is due primarily to greater accuracy on true dynamic responses.

7.1.v Conclusions

The findings from this research did not support the hypothesis of mimicry as an important mechanism underlying situated recognition of natural expressions. Whereas, the greater impact of context on judgments about dynamic expressions indicated that situated recognition relied less on motor simulation and more on alternative systems. In addition to affective and sensorimotor systems, the integration of context in judgments about dynamic systems suggested conceptual structures were activated during dynamic processing and remained active longer with longer duration of expressions. Similarly, the effect of context on longer durations of static expressions suggests conceptual structures were also activated during processing of longer presentations of static expressions but were not readily activated when static expressions were presently only briefly.

The effect of context on both static and dynamic expressions emphasizes the role of the observer in the recognition process. Recognition is not reliant solely on decoding the configural arrangements of facial features. Instead, the results of this study indicate that in a social interaction, the observer's embodied response to a shared context is an important contributing factor to understanding the emotional state of the other person.

Finally, differences in processing dynamic and static expressions and the importance of context found in this research indicate that the use of natural dynamic expressions presented in context may provide a more appropriate paradigm for research into ecologically relevant expression recognition processes.

REFERENCES

Achaibou, A., Loth, E. & Bishop, S. J. (2016). Distinct frontal and amygdala correlates of change detection for facial identity and expression. *Social Cognitive and Affective Neuroscience,11,* 225-233, DOI: 10.1093/scan/nsv104

Adelmann, P. K. & Zajonc, R. B. (1989). Facial efference and the experience of emotion. *Annual Review Psychology, 40*, 249-280

Adolphs, R. (2002). Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral & Cognitive Neuroscience Reviews,1,* 262.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences, 4,* 267-278

Ambadar, Z. S., Schooler, J. W., & Cohn, J. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science, 16*, 403-410.

Ambady, N. & Gray, H. M. (2002). On being sad and mistaken: Mood effects on the accuracy of thin-slice judgments. *Journal of Personality and Social Psychology, 83*, 947-961.

Arsalidou M., Morris, D., & Taylor, M. (2011). Converging evidence for the advantage of dynamic facial expressions. *Brain Topograhy, 24,* 149-163.

Atkinson, A. P. (2007). Face processing and empathy. In *Empathy and Mental Illness*, Eds Tom F. D. Farrow and Peter W. R. Woodruff, pp. 360-385,, Oxford:Cambridge University Press

Atkinson, A., & Adolphs, R. (2005). Visual emotion perception: Mechanisms and processes. In Lisa Feldman Barrett, Paula M. Niedenthal, Piotr Winkielman (Eds.) *Emotion and Consciousness*, pps150-182, New York: Guilford Press.

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A. & Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science, 19,* 724-732. DOI:10.1111/j.1467-9280.2008.02148.x

Back. E., Jordan, T. R. & Thomas, S. M. (2009). The recognition of mental states from dynamic and static facial expressions. *Visual Cognition, 17,* 1271-128

Balconi, M., Bortolotti, A., & Gonzaga, L. (2011). Emotional face recognition, EMG response, and medial prefrontal activity in empathic behaviour., *Neuroscience Research, 71,* 251-259.

Balconi, M. & Bortolotti, A. (2012). Resonance mechanisms in empathic behavior BEES, BIS/BAS and psychophysiological contribution, *Physiology & Behavior, 105*, 298-304.

Balconi, M. & Canavesio, Y. (2013). High-freuqency rTMS improves facial mimicry and detection responses in an empathic emotional task. *Neuroscience, 16,* 12-20. DOI: 10.1016/j.neuroscience.2012.12.059

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42,* 241-251.

Barresi, J. & Moore, C. (1996). Intentional relations and social understanding. *Behavioral and Brain Sciences, 19,* 107-154.

Barresi, J. & Moore, C. (2008). The neuroscience of social understanding. In J. Zlatev, T. Racine, C. Sinha, and E. Itkonen (Eds.) *The Shared Mind: Perspectives on Intersubjectivity*, Amsterdam/Philadelphia: John Benjamins.

Barrett, L. F., Lindquist, K. A., & Gendron, M. (2007). Language as context for the perception of emotion. *Trends in Cognitive Sciences, 11*, 327-332.

Barrett L. F., Mesquita, B. & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science, 20*, 286-290.

Barsalou. L. W. (2003). Situated simulatin the human conceptual system*. Language and Cognitive Processes, 18,* 513-562.

Barsalou, L W., (2008). Grounded Cognition. *Annual Review of Psychology. 59*, 617 645, DOI: 10.1146/annurev.psych.59.103006.093639

Bassili, J. N. (1978). Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology, 37*, 2049-2058.

Batty, M. & Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions. *Cognitive Brain Research, 17*, 613-620.

Blake, R. & Shiffrar, M. (2007). Perception of human motion. *The Annual Review of Psychology, 58,* 47-73

Bould, E., & Morris, N. (2008). Role of motion signals in recognizing subtle facial expressions of emotion. *British Journal of Psychology, 99,* 167-189, DOI: 10.1348/000712607X206702

Bould, E., Morris, N. & Wink, B. (2008). Recognizing subtle emotional expressions: The role of facial movements. *Cognition and Emotion, 22,* 1569-1587, DOI: 1.0.1080/2699930801921156

Bourgeois, P. & Hess, U., (2008). The impact of social context on mimicry. *Biological Psychology, 77,* 343-352.

Carr, L., Iacoboni, M., Dubeau, M-C., Mazziotta J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Science (Biology), 100,* 5497-5502.

Carroll, J. M. and Russell, J. A., (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology, 70,* 205-218.

Chan, L. P., Livingstone, S.R., & Russo, F. A. (2014). Facial mimicry in response to song. *Music perception, 30,* 361-367, DOI: 10.1525/mp.2013.30.4.361

Cowie, R. (2005). What are people doing when they assign everyday emotion terms? *Psychological Inquiry, 16*, 11- 48.

Cowie, R. (2009). Perceiving emotion: towards a realistic understanding of the task. Phil. Tans. R. Soc. B 364, 3515-3525. DOI: 10.1098/rstb.2009.0139.

Davis, J. I., Senghas, A., Brandt, F., & Ochsner, K. N. (2010). The effects of BOTOX injections on emotional experience. *Emotion,10*, 433-410. DOi:10.1037/a0018690

Davis, M. H. (1994/96). *Empathy; A social psychological approach*. Madison, Wis.: Brown and Benchmark Publishers.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*, 113-126.

Dimberg, U., & Thunberg, M. (2012). Empathy, emotional contagion, and rapid facial reactions to angry and happy facial expressions. *PsyCH Journal, 1*, 118-127. DOI:10.1002/pchj.4

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotion facial expressions. *Psychological Science, 11*, 86-89.

Dimberg, U., Thunberg, M., & Grunedal, S. (2002). Facial reactions to emotional stimuli: Automatically controlled emotional responses. *Cognition and Emotion, 16*, 449-471. DOI:10.1080/2699930143000356

Ekman, P. (1992). Facial expressions of emotion: New findings, new questions. *Psychological Science, 3*, 34-38.

Ekman, P. (2004). *Emotions Revealed: Recognizing faces and feelings to improve communication and emotional life*. New York: Owl Books.

Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*. New York: Pergamon Press

Ekman, P. & Friesen, W. V. (1976). Pictures of facial affect, Palo Also, CA: Consulting Psychologists Press

Ekman, P. & Friesen, W. V. (1978). Facial action coding system. Palp Alto, CA: Consulting Psychologists Press

Elfenbein, H. A. & Ambady, N. (2002). Is there an in-group advantage in emotion recognition? *Psychological Bulletin, 128*, 243-249

Enticott, P. Johnston, P J., Herring S. E. Hoy, K. E. & Fitzgerald, P. B. (2008). Mirror neuron activation is associated with facial emotion processing. *Neuropsychologia, 46*, 2851-2854

Fridlund, A. J. (1991). Sociality of solitary smiling: Potentiation by an implicit audience. *Journal of Personality and Social Psychology, 60*, 229-240.

Fiorintini, C. & Viviani, P. (2011). Is there a dynamic advantage for facial expressions? *Journal of Vision, 11,* (3), 1-15, DOI: 10.1167/11.3.17

Frith, C. D. (2007). The social brain? *Philosophical transactions of the Royal Society Biology, 362*, 671-678.

Fujimura, T., Sato, W., & Suzuki, N. (2010). Facial expression arousal level modulates facial mimicry. *International Journal of Psychophysiology*, 76, 88-92. DOI: 10.1016/j.ijpsycho.2010.02.008

Gallese, V., (2003). The roots of empathy: The shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology, 36*, 171-180

Gallese, V. (2007). Before and below ‘theory of mind’: embodied simulation and the neural correlates of social cognition. Philisophical Transactions of the Royal Society, B. 362, 659-669, DOI:10.1098/rstb.2006.2002

Gallese, V. & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2*, 493-501.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences, 8,* 396-403.

Gallese, V., & Sinigaglia, C. (2008). What is so special about embodied simulation? *Trends in Cognitive Science, 11*, 512-519, DOI:10.1016/j.tics.2011.09.003

Goldman, A. (2005). Imitation, mindreading, and simulation. In Susan Hurley and Nick Chater (Eds.) *Perspectives on Imitation, Vol. 2,* pps 79-93, Cambridge, MA: MIT Press

Goldman, A. (2009). Mirroring, simulating, and mindreading. *Mind and Language, 24*, 232-252.

Goldman, A. I. (2011). Two routes to empathy: insights from cognitive neuroscience, In Amy Coplan and Peter Goldie (Eds.), *Empathy; Philosophical & Psychological Perspectives, pp 31-44,* Oxford Publishers Online.

Goldman, A. & Sripada, C. (2005). Simulationist models of face-based emotion recognition. *Cognition, 94*, 151-161.

Gopnik, A. & Wellman, The theory-theory. In: Hirschfeld, L, Gelman S. (Eds.), *Mapping the Mind*. pp 257-293 Cambridge University Press, NewYork, .

Halberstadt, J., Winkielman, P., Niedenthal, P. M., & Dalle, N. (2009). Emotional conception: How embodied emotion concepts guide perception and facial action. *Psychological Science, 20*, 1254-61.

Hassin, R.R., Aviezer, H. A., & Bentin, S. (2013). Inherently ambiguous: Facial expressions of emotions, in context. *Emotion Review, 5*, 60-65.

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1992). Primitive emotional contagion. *Review of Personal and Social Psychology, 14,* 151-177.

Haxby, J. V.. Hoffman, E. A. & Gobbini, M.I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Science, 4*, 223-233

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry, 51,* 59-67.

Heberlein, A. & Atkinson, A. (2009). Neuroscientific evidence for simulation and shared substrates in emotion recognition: Beyond faces. *Emotion Review. 1,* 162-177. DOI:10.1177/175073908100441

Hennelotter, A. Schroeder, U., Erhard, P, & Ceballos-Baumann, A. (2005). A common neural basis for receptive and expressive communication of pleasant facial affect. *Neuroimage, 26*, 581-591. DOI: 10.1016/j.neuroimage.2005.01.057

Hoffman , H. Traue, H., Limbrecht-Ecklundt, K., Walter, S. & Kessler, H. (2013. Static and dynamic presentation of emotions in different facial areas: Fear and surprise show influences of temporal and spatial properties., *Psychology, 4,*663-668.

Hess, U., & Blairy, S. (2001). Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *International Journal of Psychophysiology*, *40*, 129-41.

Hess, U. & Fischer, A. (2014). Emotional mimicry: Why and when we mimic emotions. *Social and Personality Psychology Compass, 8,* 45-57, DOI:10.1111/spc3.12083

Iacoboni, M. (2009). Imitation, empathy, and mirror neurons, *Annual Review of Psychology, 60,* 653-70, DOI:10.110707.163604

Iacoboni, M. (2008). Mirroring People, New York: Farrar, Strauss, and Giroux.

Kilts, C.D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *NeuroImage, 18*, 156-168.

Kohler, C.G. et al. (2003). Facial emotion recognition in schizophrenia: intensity effects and error pattern. *American Journal of Psychiatry, 10*, 1768-74.

Kret, M. E. (2015). Emotional expressions beyond facial muscle actions. A call for studying autonomic signals and their impact on social perception. *Frontiers in Psychology,REVIEW,* 27 May 2015, DOI: 10.3380/fpsyg.2015.00711

Krumhuber. E. G., Kappas, A., & Manstead, A. S. R. (2013). Effects of dynamic aspects of facial expressions: A review. *Emotion Review, 5,* 41-46. DOI: 10.1177/1754073912451349

LaBar, K. S., Crupain, M. J., Voyvodic, J. T. & McCarthy, G. (2003). Dynamic perception of facial affect and identity in the human brain. *Cerebral Cortex, 13*, 1023-1033

Laird, J. D. (1974). Self-attribution of emotion: The effects of expressive behavior on the quality of emotional experience. *Journal of Personality and Social Psychology, 29*, 475-486.

Lakin, J. L., Jefferis, V. E., Cheng, C. M. & Chartrand, T. (2003). The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. *Journal of Nonverbal Behavior, 23,* 145-162.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture System (IAPS): Affective ratings of pictures and instruction manual*. Technical Report 1-6, University of Florida, Gainesville, FL.

Livingstone, S. T., Esztella, V., McGarry, L. M., Lang, A. E. & Russo, F. A. (2016). Deficits in the mimicry of facial expressions in Parkinson's Disease. *Frontiers in Psychology, 7:780*. DOI: 10.3389/fpsyg.2016.00780

Maringer, M., Krumhuber, E. G., Fischer, A. H., & Niedenthal, P. M. (2011). Beyond Smile Dynamics: Mimicry and beliefs in Judgments of smiles. *Emotion, 11*, 181-187, DOI: 10.1037/a0022596

Masuda, T., Ellsworth, P.C., Mesquita, B., Leu, J., Tanida, S. & van de Veerdonk, E. (2008). Placing the face in context: Cultural differences in the perception of facial emotion. *Journal of Personality and Social Psychology, 94, 365-381.*

Matsumoto, D., & Hwang, H. S. (2010). Judging faces in context. *Social and Personality Psychology Compass, 4/6*, 393-402.

McRorie, M. & Sneddon, I. (2007). Real emotion is dynamic and interactive. *Affective Computing and Intelligent Interaction, 4738,* 759-760. DOI: 10.1007/978/-3-540-748889-2_86

Meeren, K. K. M., van Heijnsbergen, C.C.R.J., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Science, 102,* 16518-16523.

Meier, B. P., Schnall, S., Schwarz, N. & Bargh, J. A. (2012). Embodiment in social psychology. *Topics in Cognitive Science, 4,* 705-716, DOI:10.1111/j.1756 8765.2012.01212.x

Mikels, J.A., Fredrickson, B. L., Larkin, G. R. & Lindberg, C. M. (2005). Emotional category data on images from the International Affective Picture System. *Behavior Research Methods, 37*, 626-630.

Moody, E. J., McIntosh, D. N., Mann, L. J., & Weisser, K. R. (2007). More than mere mimicry? The influence of emotion on rapid facial reaction to faces. *Emotion, 7*, 447-457. DOI: 10.1037/1528-3542.7.2.447.

Neal, D. T. & Chartrand, T. L. (2011). Embodied emotion perception: Amplifying and dampening facial feedback modulates emotion perception accuracy. *Social Psychological and Personality Science, 2,* 673-678.

Niedenthal, P. M. (2007). Embodying emotion. *Science, 316*,1002-1005.

Niedenthal, P. M., Barsalou, L.W. ., Ric F. & Krauth-Gruber, S. (2005). Embodiment in the acquisition and use of emotional knowledge. In *Emotion and Consciousness,* (Eds.) *L. M. Barret, P.M. Niedenthal.,* Winkielman, P. Guilford Press:New York

Niedenthal, P. M., Barsalou, L. W., Winkielman, P., Krauth-Gruber, S., & Ric F. (2005). Embodiment in attitudes, social perception, and emotion. *Personality and Social Psychology Review, 9*, 184-211.

Niedenthal, P. M., Brauer, M., Halberstadt, J. B., & Innes-Ker, A. H. (2001). When did her smile drop? Facial mimicry and the influences of emotional state on the detection of change in emotional expression. *Cognition and Emotion, 15*, 853-864.

Niedenthal, P. M., Halberstadt, J. B., Margolin, J, & Innes-Ker, A. H. (2000). Emotional state and the detection of change in facial expression of emotion. *European Journal of Social Psychology, 30*, 211-222.

Niedenthal, P. M., Mermillod, M., Maringer, M., & Hess, U. (2010). The simulation of Smiles (SIMS) model: Embodied simulation and the meaning of facial expression. *Behavioral and Brain Sciences, 33*, 6, 447-463.

Niedenthal, P. M. & Maringer, M. (2009). Embodied emotion considered. *Emotion Review, 1,* 122-128.

Oberman, L. M., Winkielman, P., & Ramachandran, V. S. (2007). Face to face: Blocking facial mimicry can selectively impair recognition of emotional expressions. *Social Neuroscience, 2,* 167-178.

Ochsner, K., Knierim, K., Ludlow, D., Hanelin, J., Ramachandran, T., Govler, G., & Mackey, S., (2004). Reflecting upon feelings: An fMRI study of neural systems supporting the attribution of emotion to self and other. *Journal of Cognitive Neuroscience, 16,* 1746-1772.

Pelphry, K. A. & Carter, E. (2008). Charting the typical and atypical development of the social brain. *Developmental and Psychopathology, 20,* 1081-1102. DOI: 10.1017/S0954579408000515

Pitcher D., Dilks, D. D., Saxe,R. R. Triantafyllou, C. & Kanwisher, N. (2011). Differential selectivity for dynamicversus static information in face-selective regions. *NeuroImage, 56*, 2356-63. DOI:10.1016/j.neuroimage.2011.03.067

Ponari M, Conson M, D'Amico NP, Grossi D, Trojano L (2012) Mapping correspondence between facial mimicry and emotion recognition in healthy subjects. *Emotion, 12*: 1398–1403. doi: 10.1037/a0028588

Preston, S. D. & de Waal, F. W. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25*, 1-72,

Prinz. W. (1997). Perception and action planning*, European Journal of Cognitive Psychology, 9*, 129-154.

Prinz, W. (2005). An ideomotor approach to imitation. In Susan Hurley and Nick Chater, (Eds.) *Perspectives in Imitation, Vo.l 1*, pp 141-156. Cambridge, MA: MIT Press.

Righart, R. & de Gelder, B. (2006). Context influences early perceptual analysis of faces. An electrophysiological study. *Cerebral Cortex, 16*, 1249-1257.

Righart, R. & de Gelder, B. (2008). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive, Affective, & Behavioral Neuroscience, 8*, 264 272, DOI: 10.3758/CABN.8.3.264

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L., (1996). Premotor cortex and the recognition of motor actions. *Cognition and Brain Research, 3,* 131-141.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin, 11*, 102-141.

Russell, J. A., Bachorowski, J. & Fernandez-Dols, J-M. (2003). Facial and vocal expressions of emotion. *Annual Review Psychology, 54*, 329-49.

Rychlowska, M., Canadas, E., Wood, A., Krumhuber, E., Fischer, A. & Niedenthal, P.M. (2014). Blocking mimicry makes true and false smiles look the same. *PLoS ONE, 9* (3),e90876. DOI:10.1371/journal.pone.00908076

Rymarczyk, K., Biele, C., Grabowska, A., & Majczynski, H. (2011). EMG activity in response to static and dynamic facial expressions*. International Journal of Psychophysiology, 79,* 330-333

Sato, W., Fujimara & Suzuki (2008). Enhance facial EMG activity in response to dynamic facial expressions. *International Journal Psychophysiology, 70*, 70-74.

Sato, W., & Yoshikawa, S. (2004). The dynamic aspects of emotional facial expressions. *Cognition and Emotion, 18*, 701-710

Sato, W., & Yoshikawa, S. (2007). Spontaneous facial mimicry in response to dynamic facial expressions. *Cognition, 104,* 1-18.

Saxe, R. (2005). Against simulation: The argument from error. *Trends in Cognitive Sciences, 9,* 174-179.

Schulz, J. & Pilz, K. S. (2009). Natural facial motion enhances cortical responses to faces. *Experimental Brain Research, 194,* 465-475. DOI: 10.1007/s00221-009 1721-9

Sebanz, N, Knoblich, G., & Prinz, W. (2003). Representing others' actions: just like one's own? *Cognition, 88*, B11-B21

Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: a double dissociation emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain, 132,* 617-627. DOI:10.1093/brain/awn279

Sonnby-Borgström, M. (2002). Automatic mimicry reactions as related to differences in emotional empathy, *Scandinavian Journal of Psychology, 43*, 433-443.

Spunt, R. P., & Lieberman, M. D. (2013). The busy social brain: Evidence for Automaticity and control in the neural systems supporting social cognition and action understanding. *Psychological Science, 24,* 80-6. DOI: 10.1177/0956797612450884

Strack. F., Martin, L.L., Stepper, S. (1988) Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of facial feedback hypothesis. *Journal of Personality and Social Psychology, 54,* 768-777

Susskind, J.M., Littlewort, G., Bartlett, M.S., Movellan, J., & Anderson, A.K. (2007). Human and computer recognition of facial expressions of emotion. *Neuropsychologia, 45*, 487-494.

Thornton, I. M. & Kourtzi, Z. (2002). A matching advantage for dynamic faces. *Perception, 31*, 113-132

Trautmann, S. A., Fehr, T., & Herrmann, M. (2009). Emotions in motion: Dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion –specific activations. *Brain Research, 1284*, 100-115

Trautmann, S. A., Dominguez-Borras,J., Escera, C. Herrmann, M., & Fehr, T. (2013). The perception of dynamic and static facial expressions on happiness and disgust investigated by ERP and fMRI constrained source analysis. *PlosOne*, June 20, 2013, DOI:org/10.1371/journal:pone.0066997.

Van Overwalle, F. & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta analysis. *Neuroimage, 48*, 564-84. DOI: 10.1016/j.neuroimage.2009.06.009

Weyers, P., Muhlgerger A., Hefele, C., & Pauli, P. (2007). Electromyographic responses to static and dynamic avatar emotional facial expressions. *Psychophysiology, 43*, 450-3.

Wild, B., Erb, M., & Bartels, M. (2001). Are emotions contagious? Evoked emotions while viewing emotionally expressive faces: quality, quantity, time course, and gender differences. *Psychiatry Research, 102*, 109-124.

Winkielman, P., Niedenthal, P. M., & Oberman, L. (2008). The embodied emotional mind. In G. R. Semin & E. R. Smith (Eds.), *Embodied Grounding; Social, Cognitive, Affective and Neuroscientific Approaches*, pp 263-288, New York: Cambridge University Press.

Zajonc, R. B., & Markus, H., (1984). Affect and cognition: the hard interface. In Carroll E. Izard, Jerome Kagan, Robert B. Zajonc (Eds.) *Emotions, Cognition, & Behavior,* pp. 73-102.

Zaki, J. (2013). Cue integration: A common framework for social cognition and physical perception. *Perspectives on Psychological Science, 8,* 296-312.

APPENDIX 1

TABLE A1. ORGANIZATION OF CONTEXT STIMULI

Table 1. Organization of Paired Context Images

Valence	IAPS#	Image	Set	Pair	Category *
Different	5260	Waterfalls	1	1	AwCE
	1280	Rat			DF
Different	2340	Grampa & kids	1	2	AmCE
	9520	Dirty boys by water			S
Same +	1850	Camels	1	3	Aw
	1811	Chimps			Am
Same -	2691	Man throwing rock	1	4	U
	6212	Gun at child			U
Different	2058	Baby	2	1	C
	6550	Knife			U
Different	5621	Sky Dive	2	2	AwE
	6838	Child screaming			S
Same +	2091	Kittens	2	3	AmC
	8030	Ski			E
Same -	2205	Death Bed	2	4	S
	9042	Rod through lip			D
Different	1340	Parrots	3	1	Am
	9810	Klu Klux Klan			ADS
Different	9300	Toilet	3	2	D
	2655	Dog and baby			AmC
Same +	5950	Lightning	3	3	AwE
	8117	Goalie			E
Same -	2900	Boy crying	3	4	S
	9410	Dead child			S
Different	5600	Mountain	4	1	Am
	3530	Gun at man			U
Different	1932	Shark	4	2	F
	8180	Water dive			AmAwE
Same +	2310	Mother & infant	4	3	C
	5910	Fireworks			AwE
Same -	1300	Rabid dog	4	4	DF
	9530	Boys & industrial waste			S

Note: * Normed ratings by Mikels, et al. (2005). Am = Amusement, Aw = Awe, C = Content, E = Excitement, A = Anger, D = Disgust, F = Fear, S = Sad, U = Undifferentiated (similar ratings across categories).

APPENDIX 2

EARLY USE OF THE PARADIGM

EARLY USE OF THE PARADIGM

The paradigm was initially applied in a study testing for an effect of familiarity on emotion recognition. During this earlier study, we were also able to test programming parameters and ascertain that participants could discriminate complex and possibly ambiguous emotion expressions (eg. blended emotions, and expressions that were presented with same-valence foils) in naturalistic facial displays.

Sixty participants (*N* = 60, mean age = 21 yrs) observed 5 s dynamic facial expressions of 2 male targets in response to 32 context images representing 8 emotions. In a single trial a 5 s facial expression was presented simultaneously with one context image to which the expression was the true response (the match), or not (the mismatch). As in subsequent studies, context images were organized as pairs so that the foil would be the either the same or different valence. Prior to testing, half the participants experienced a familiarity manipulation with one target, half with the other target. During this manipulation they observed two presentations of one target's true responses to 16 of the context images. During testing participants were exposed to 2 blocks of 64 trials (2 targets, 32 images, 2match/mismatch). The dependent variable was recognition accuracy, defined as the proportion of responses correctly identifying the match and mismatch.

Data were submitted to a mixed ANOVA with familiar target as the between subjects factor, and target, old/new images, match/mismatch, and same/different valence as within-subject factors. Data were distributed normally, without ceiling or floor effects. The mean proportion of correct responses, *M* = .66, *SD* = .05, was comparable to emotion recognition studies measuring accuracy on normal populations [ranging from *M* = .62.5 (Hess, & Blairy, 2001) to *M* = .71 (Oberman, et al., 2007) averaged across four, positive and negative emotions, and *M* = .76 for typical females tested with the

BMET (Neal & Chartrand, 2011)]. We observed no significant change in accuracy over blocks, $p = .85$, and so ruled out the likelihood of practice effects. Mimicry was not yet a factor of interest in this study.

Participants showed better accuracy with the familiar face, $M = .73$, $SD = .08$, relative to the unfamiliar face, $M = .63$, $SD = .08$, when presented with familiar images (exposed to during the familiarity manipulation) but the advantage did not generalize to new images: The two variable of familiar face and familiar images interacted so that in response to new contextual images, participants were no more accurate with the familiar face, $M = .63$, $SD = .08$, than with the unfamiliar face, $M = .63$, $SD = .08$, , $p = < .001$.

Regarding complexity of stimuli: As anticipated, participants were more accurate in their recognition of a facial expression as the true response to an image when context pairs were of a different valence relative to same-valence context pairs, $M_{Different} = .73$, $SD = .07$ $M_{Same} = .58$, $SD = .06$, $p < .0001$. Participants were better able to judge the face as the match than the mismatch, $M_{Match} = .73$, $SD = .11$, $M_{Mismatch} = .59$, $SD = .11$, $p < .001$. They showed no difference in accuracy at judging the face as the match and mismatch on different valence trials, $M_{Match} = .74$, $SD = 10$, $M_{Mismatch} = .73$, $SD = .13$, and the match on same valence trials, $M_{Match} = .74$, $SD = 10$, but their judgment of the face as the mismatch on same valence trials was below chance, $M_{Mismatch} = .44$, $SD = .14$, $p < .0001$. Criterion *c* analysis confirmed a bias to judge the face as a match on these more complex trials involving context pairs of the same valence, $p = .01$.

Familiarity interacted with same/different-valence context pairs, $p = .004$: On difficult trials in which the face was presented with a foil image of the same valence, participants were significantly more accurate on the familiar face than the unfamiliar face. There was no difference due to familiarity on the easier trials involving different valence context pairs.

From this earlier study we determined we could apply the paradigm to effectively measure recognition accuracy without encountering either ceiling or floor effects. We considered the response bias to judge the face as the match and the interaction with familiarity and same or different valence trials as support for our theoretical model for the paradigm: That observers would likely use their own response as a reference when making a judgment about the emotional response of another individual, particularly in more complex situations – when the person is unfamiliar and when the emotion expressions are more complex and possibly ambivalent.

APPENDIX 3

THE INTERACTION OF PRESENTATION ORDER, PEN 1ST, AND DISRUPTED MIMICRY IN EXPERIMENT 1

The Interaction of Presentation Order, Pen 1^{st}, and Disrupted Mimicry in Experiment 1

	Presentation Order	Pen1st	Mimicry	N	Mean	SD	FLSD	Block	
1	1	Pen 2nd	NoPen	10	0.7179688	0.06341026	0.02988485	1,3	
2	1	Pen 2nd	Pen	10	0.6703125	0.05570726	0.02988485	2,4	
3	1	Pen 1st	NoPen	10	0.7046875	0.04626337	0.02988485	2,4	
4	1	Pen 1st	Pen	10	0.7234375	0.04573261	0.02988485	1,3	Order 1234
5	2	Pen 2nd	NoPen	10	0.6773438	0.03845888	0.02988485	1,3	
6	2	Pen 2nd	Pen	10	0.7234375	0.04573261	0.02988485	2,4	Order 2341
7	2	Pen 1st	NoPen	10	0.7343750	0.05645699	0.02988485	2,4	
8	2	Pen 1st	Pen	10	0.6585937	0.04085306	0.02988485	1,3	
9	3	Pen 2nd	NoPen	8	0.7285156	0.03360283	0.02988485	1,3	Order 3412
10	3	Pen 2nd	Pen	8	0.6865234	0.05155870	0.02988485	2,4	
11	3	Pen 1st	NoPen	7	0.6808036	0.05871144	0.02988485	2,4	Less difference for Pen1st Group
12	3	Pen 1st	Pen	7	0.6997768	0.06672825	0.02988485	1,3	
13	4	Pen 2nd	NoPen	8	0.6660156	0.05392485	0.02988485	1,3	Order 4123
14	4	Pen 2nd	Pen	8	0.7167969	0.04149775	0.02988485	2,4	
15	4	Pen 1st	NoPen	8	0.7304688	0.03340766	0.02988485	2,4	
16	4	Pen 1st	Pen	8	0.6904297	0.05062002	0.02988485	1,3	

The 3-way interaction of Presentation Order, Pen 1^{st} and Disrupted Mimicry, ($p < .0008$) is explained by participants consistently showed better accuracy during blocks 1 and 3, regardless of experiencing the pen manipulation 1^{st} or 2^{nd} and when mimicry was disrupted mimicry or possible. The only group who did not show the pattern of better accuracy on blocks 1 and 3 were participants who experienced a pen manipulation first and were presented with the 3^{rd} order of presentation, which meant they saw block 3 first (Blocks 3,4,1,2). To better understand how the interaction affected accuracy required an analysis of the data with additional variables, such as splitting the data by Time so that it could be determined if the difference in accuracy was in Block 1 or 3. There were no other significant higher order interactions with these variables and therefore further investigation was not done.

APPENDIX 4

ANOVA TABLES FOR EXPERIMENT 4

ANOVA tables for Experiment 1 reported in Chapter 3

Main Model

Bonferroni correction .05 / 15 tests = criterion of .003

$ANOVA

	Effect	DFn	DFd	F	p	p<.0033	ges
2	hold.bite	1	69	1.56898187	2.145837e-01		2.987386e-03
3	Pen.vs.NoPen	1	69	1.76474303	1.884111e-01		1.818544e-03
5	ModelResponse	1	69	0.51999519	4.732800e-01		3.903718e-03
7	Congruence	1	69	450.92284633	5.640671e-32	*	3.288621e-01
4	hold.bite:Pen.vs.NoPen	1	69	0.20581424	6.514918e-01		2.124300e-04
6	hold.bite:ModelResponse	1	69	0.10695443	7.446285e-01		8.054278e-04
8	hold.bite:Congruence	1	69	2.65335276	1.078893e-01		2.875043e-03
9	Pen.vs.NoPen:ModelResponse	1	69	0.07956347	7.787342e-01		5.757636e-05
11	Pen.vs.NoPen:Congruence	1	69	0.14266053	7.068097e-01		8.996015e-05
13	ModelResponse:Congruence	1	69	223.74451517	2.420329e-23	*	1.890296e-01
10	hold.bite:Pen.vs.NoPen:ModelResponse	1	69	1.53301987	2.198560e-01		1.108209e-03
12	hold.bite:Pen.vs.NoPen:Congruence	1	69	0.03771022	8.465968e-01		2.378122e-05
14	hold.bite:ModelResponse:Congruence	1	69	0.04030601	8.414741e-01		4.198788e-05
15	Pen.vs.NoPen:ModelResponse:Congruence	1	69	1.23462410	2.703665e-01		6.554386e-04
16	hold.bite:Pen.vs.NoPen:ModelResponse:Congruence	1	69	0.22870970	6.339940e-01		1.214825e-04

Select analyses

BITE group Bonferroni correction .05/7=.007

$ANOVA

	Effect	DFn	DFd	F	p	p<.007	ges
2	Pen.vs.NoPen	1	35	1.54096906	2.227291e-01		3.503446e-03
3	ModelResponse	1	35	0.09145707	7.641229e-01		1.291820e-03
4	Congruence	1	35	353.48838115	7.132241e-20	*	3.797613e-01
5	Pen.vs.NoPen:ModelResponse	1	35	0.38883555	5.369551e-01		6.883170e-04
6	Pen.vs.NoPen:Congruence	1	35	0.01723425	8.963061e-01		2.362054e-05
7	ModelResponse:Congruence	1	35	122.15634610	5.802170e-13	*	2.053667e-01
8	Pen.vs.NoPen:ModelResponse:Congruence	1	35	1.19181946	2.824211e-01		1.435045e-03

HOLD group Bonferroni correction .05/7=.007

$ANOVA

	Effect	DFn	DFd	F	p	p<.007	ges
2	Pen.vs.NoPen	1	34	0.3866458	5.382153e-01		0.0007227812
3	ModelResponse	1	34	0.4867622	4.901192e-01		0.0077021878
4	Congruence	1	34	150.1480472	5.034304e-14	*	0.2791394212
5	Pen.vs.NoPen:ModelResponse	1	34	1.3772510	2.487257e-01		0.0015877118
6	Pen.vs.NoPen:Congruence	1	34	0.1662682	6.860062e-01		0.0001934743
7	ModelResponse:Congruence	1	34	102.4074985	8.594693e-12	*	0.1740549602
8	Pen.vs.NoPen:ModelResponse:Congruence	1	34	0.2072609	6.518168e-01		0.0001934743

APPENDIX 5

ANOVA TABLES FOR EXPERIMENT 2

ANOVA tables for Experiment 2 reported in Chapter 5

ANOVA for Main Model
Bonferroni correction for multiple comparisons: .05 / 15 = .003

	Effect	DFn	DFd	F	p	p<.05	ges	
2	Group	1	93	7.779358e+00	6.409980e-03	*	4.728982e-03	NS
3	Valence	1	93	1.340767e+02	1.000582e-19	*	5.664893e-02	
5	ModelResponse	1	93	1.672952e+01	9.158982e-05	*	8.120933e-02	
7	Pen	1	93	1.296486e+00	2.577808e-01		8.949742e-04	
4	Group:Valence	1	93	5.851531e-03	9.391894e-01		2.620796e-06	
6	Group:ModelResponse	1	93	1.689387e+00	1.968942e-01		8.846594e-03	
8	Group:Pen	1	93	1.411984e-02	9.056688e-01		9.755670e-06	
9	Valence:ModelResponse	1	93	8.141939e-02	7.760176e-01		9.952075e-05	
11	Valence:Pen	1	93	3.685573e-01	5.452711e-01		2.963960e-04	
13	ModelResponse:Pen	1	93	5.225878e-02	8.196804e-01		5.244676e-05	
10	Group:Valence:ModelResponse	1	93	2.784363e+00	9.855197e-02		3.392182e-03	
12	Group:Valence:Pen	1	93	6.799215e-01	4.117237e-01		5.466601e-04	
14	Group:ModelResponse:Pen	1	93	1.096829e-01	7.412506e-01		1.100711e-04	
15	Valence:ModelResponse:Pen	1	93	2.109643e+00	1.497391e-01		1.451970e-03	
16	Group:Valence:ModelResponse:Pen	1	93	2.457095e-04	9.875272e-01		1.693564e-07	

Select ANOVA for the Dynamic group

Bonferroni Correction .05/7=.007

	Effect	DFn	DFd	F	p	p<.05	ges
2	Valence	1	46	87.33631142	3.350018e-12	*	6.424968e-02
3	ModelResponse	1	46	17.89625600	1.099556e-04	*	1.482551e-01
4	Pen	1	46	0.67979188	4.139134e-01		1.233574e-03
5	Valence:ModelResponse	1	46	2.32699871	1.339935e-01		5.333041e-03
6	Valence:Pen	1	46	1.21323856	2.764219e-01		1.879540e-03
7	ModelResponse:Pen	1	46	0.00441615	9.473043e-01		1.274422e-05
8	Valence:ModelResponse:Pen	1	46	1.08022992	3.040785e-01		1.602434e-03

Select ANOVA for the Static group

Bonferroni Correction .05/7 = .007

	Effect	DFn	DFd	F	p	p<.05	ges
2	Valence	1	47	54.56208545	2.124567e-09	*	0.0506311669
3	ModelResponse	1	47	3.34207473	7.388032e-02		0.0360322849
4	Pen	1	47	0.61957912	4.351552e-01		0.0006477480
5	Valence:ModelResponse	1	47	0.80572657	3.739620e-01		0.0020478185
6	Valence:Pen	1	47	0.01915256	8.905212e-01		0.0000315727
7	ModelResponse:Pen	1	47	0.20997457	6.488991e-01		0.0002794268
8	Valence:ModelResponse:Pen	1	47	1.03342619	3.145600e-01		0.0013347396

APPENDIX 6

ANOVA TABLES FOR EXPERIMENT 3

ANOVA Tables for Experiment 3, reported in Chapter 6

BONFERRONI CORRECTION: .05/31 = .0016

$ANOVA

	Effect	DFn	DFd	F	p	p<.05	ges
2	staticdynamic	1	59	2.061543e+01	2.820527e-05	*	2.890108e-02
3	pen	1	59	1.476158e+00	2.292166e-01		2.126506e-03
5	expected_response	1	59	8.895843e+00	4.150645e-03	*	6.208829e-02
9	valence	1	59	2.300547e+02	5.054330e-22	*	2.387734e-01
13	video_time	1	59	6.288361e+00	1.492492e-02	*	1.660524e-02
4	staticdynamic:pen	1	59	1.882138e+00	1.752856e-01		2.709763e-03
6	staticdynamic:expected_response	1	59	8.560056e+00	4.871316e-03	*	5.988502e-02
7	pen:expected_response	1	59	1.026994e+00	3.150038e-01		7.584415e-03
10	staticdynamic:valence	1	59	1.984126e+00	1.642047e-01		2.697967e-03
11	pen:valence	1	59	8.318462e-02	7.740381e-01		1.134056e-04
14	staticdynamic:video_time	1	59	8.463599e+00	5.101778e-03	*	2.222160e-02
15	pen:video_time	1	59	3.841440e-01	5.377801e-01		1.030448e-03
17	expected_response:valence	1	59	3.540380e+02	1.304353e-26	*	2.328718e-01
21	expected_response:video_time	1	59	7.875050e+00	6.780177e-03	*	8.440204e-03
25	valence:video_time	1	59	1.529906e-01	6.971027e-01		1.688278e-04
8	staticdynamic:pen:expected_response	1	59	2.273269e-01	6.352743e-01		1.688797e-03
12	staticdynamic:pen:valence	1	59	3.465424e-03	9.532563e-01		4.724927e-06
16	staticdynamic:pen:video_time	1	59	1.044870e-01	7.476531e-01		2.804918e-04
18	staticdynamic:expected_response:valence	1	59	6.196787e+00	1.563750e-02	*	5.285236e-03
19	pen:expected_response:valence	1	59	5.157314e+00	2.681304e-02	*	4.402573e-03
22	staticdynamic:expected_response:video_time	1	59	5.109600e-01	4.775425e-01		5.519857e-04
23	pen:expected_response:video_time	1	59	2.707422e-01	6.047815e-01		2.925564e-04
26	staticdynamic:valence:video_time	1	59	1.939583e+00	1.689400e-01		2.136153e-03
27	pen:valence:video_time	1	59	1.231872e+01	8.657211e-04	*	1.341385e-02
29	expected_response:valence:video_time	1	59	2.931913e-01	5.902216e-01		2.852858e-04
20	staticdynamic:pen:expected_response:valence	1	59	1.773780e+00	1.880373e-01		1.518585e-03
24	staticdynamic:pen:expected_response:video_time	1	59	1.782997e-02	8.942295e-01		1.927183e-05
28	staticdynamic:pen:valence:video_time	1	59	1.488556e+00	2.272950e-01		1.640231e-03
30	staticdynamic:expected_response:valence:video_time	1	59	2.961574e+00	9.050561e-02		2.874256e-03
31	pen:expected_response:valence:video_time	1	59	8.194502e-02	7.756814e-01		7.975189e-05
32	staticdynamic:pen:expected_response:valence:video_time	1	59	1.411462e-01	7.084916e-01		1.373607e-04

Select ANOVA for the DYNAMIC Group Bonferroni Correction .05/15 =.003

$ANOVA

	Effect	DFn	DFd	F	p	p<.05	ges	
2	pen	1	30	8.154390e-03	9.286477e-01		3.227368e-05	
3	modelresponse	1	30	8.937216e-03	9.253113e-01		1.290822e-04	
5	congruence	1	30	1.088019e+02	1.697689e-11	*	3.154863e-01	
7	video_time	1	30	3.795847e+01	8.896628e-07	*	8.747800e-02	
4	pen:modelresponse	1	30	1.396440e+00	2.466034e-01		1.977285e-02	
6	pen:congruence	1	30	4.761905e-02	8.287370e-01		2.016763e-04	
8	pen:video_time	1	30	1.153355e+00	2.914105e-01		2.904334e-03	
9	modelresponse:congruence	1	30	2.450166e+02	5.618634e-16	*	3.226986e-01	1
11	modelresponse:video_time	1	30	7.750702e+00	9.205754e-03	*	1.607640e-02	NS
13	congruence:video_time	1	30	7.526132e-01	3.925368e-01		1.160542e-03	
10	pen:modelresponse:congruence	1	30	5.020747e-01	4.840599e-01		9.753581e-04	
12	pen:modelresponse:video_time	1	30	9.568768e-02	7.592056e-01		2.016763e-04	
14	pen:congruence:video_time	1	30	4.703833e+00	3.815148e-02	*	7.209459e-03	NS
15	modelresponse:congruence:video_time	1	30	6.148055e-01	4.391317e-01		1.578964e-03	
16	pen:modelresponse:congruence:video_time	1	30	2.007528e-01	6.573330e-01		5.161290e-04	

Select Anova for the Static Group Bonferroni Correction .05/15 =.003

$ANOVA

	Effect	DFn	DFd	F	p	p<.05	ges	
2	pen	1	29	3.867535e+00	5.886409e-02		8.155608e-03	
3	modelresponse	1	29	1.437750e+01	7.020227e-04	*	1.800916e-01	3
5	congruence	1	29	1.306229e+02	2.933359e-12	*	1.770515e-01	1
7	video_time	1	29	6.336512e-02	8.030295e-01		4.725028e-04	
4	pen:modelresponse	1	29	1.138980e-01	7.381799e-01		1.737024e-03	
6	pen:congruence	1	29	3.617563e-02	8.504784e-01		5.957964e-05	
8	pen:video_time	1	29	2.552940e-02	8.741627e-01		1.904220e-04	
9	modelresponse:congruence	1	29	1.227814e+02	6.121963e-12	*	1.605085e-01	2
11	modelresponse:video_time	1	29	1.783670e+00	1.920877e-01		3.929717e-03	
13	congruence:video_time	1	29	1.046249e+00	3.148298e-01		2.816728e-03	
10	pen:modelresponse:congruence	1	29	6.051219e+00	2.009652e-02	*	9.335098e-03	NS
12	pen:modelresponse:video_time	1	29	1.759290e-01	6.779863e-01		3.889784e-04	
14	pen:congruence:video_time	1	29	7.714920e+00	9.501003e-03	*	2.040390e-02	NS
15	modelresponse:congruence:video_time	1	29	2.830415e+00	1.032330e-01		4.236511e-03	
16	pen:modelresponse:congruence:video_time	1	29	4.954358e-03	9.443686e-01		7.447084e-06	