Facial Expressions and Context Effects

A Thesis Presented

in Partial Fulfillment of the Requirements

for the Degree of of Doctor of Philosophy

in Psychology

at Massey University, Albany,

New Zealand

Hui Xu

2016

Abstract

It is common and important for us to recognise facial expressions in our daily life. Research on recognition of facial expressions was often carried out using isolated faces, which leads us to ignore accompanied contextual information (e.g. vocal sound, body language). Chapter 4 used bodily and vocal expressions as contextual stimuli to investigate whether there are context effects on recognition of all six basic facial expressions. The results generally showed that recognition of facial expressions benefits from congruent contextual stimuli, while recognition of facial expressions is impaired by incongruent contextual stimuli. Chapter 5 examined whether the observed context effects vary with the level of intensity of facial expressions. The results showed that context effects are influenced by the level of intensity of facial expressions and revealed the opposite trend of the magnitude of facilitation effects and interference effects as level of intensity of facial expressions was increased. The following chapter 6 investigated another important aspect of context effects, that is, whether attentional resources influence the observed context effects. The results showed that the magnitude of context effects was reduced when the perceptual load of task-relevant tasks was increased, at least for context effects from bodily expressions to the recognition of disgusted facial expressions. All the data collected showed commonalities and differences in the pattern of context effects on recognition of facial expressions. Future studies might concentrate on the differences among these facial

expressions to explore whether there exists a consistent pattern of context effects for all six facial expressions or to refine the existing models regarding recognition of facial expressions to better predict context effects for facial expression recognition.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my primary supervisor Dr. Heather Buttle for the continuous support of my Ph.D. study and related research, for her patience, motivation, and immense knowledge. Her guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D. study.

Besides my primary supervisor, I would like to thank my co-supervisor Dr. Stephen Hill, for his insightful comments and encouragement.

My sincere thanks go to my husband who gave me support financially, spiritually through the whole period of through all my Ph.D. study. Without him, I could not complete the whole PhD study.

Last but not the least, I would like to thank my parents for supporting me spiritually throughout writing this thesis and my life in general.

Contents

List of Figures

List of Tables

Chapter 1 Introduction

Faces can provide important cues for communication in social life. Glancing at other people's faces, we can acquire information about identity, age, sex, race, and emotion. It is of great significance for us to understand other's emotion. Knowing emotions of others gives us clues about whether we can approach (when they are happy) or avoid them (when they are angry) and about which action they will take. A large amount of research has been carried out on emotion recognition from facial expressions.

Dating back to 19th century, Darwin elaborated a chief point – facial expressions are common to humans and animals in his book *The Expression of Emotions in Man and Animals* (Darwin, 1872). He carried out cross-cultural studies, revealing that facial expressions can be recognised universally. His idea was also demonstrated by other researchers (Tomkins, 1962; Izard, 1971; Ekman, 1972) later on. Moreover, after doing research on many different cultural groups, even in preliterate groups, Ekman (1972) proposed there are six basic emotions that are delivered and expressed by different facial configurations. Based on Ekman's work, a number of databases of emotional facial expressions were built for helping researchers to explore the secrets of emotion. So far, a great deal of research has been carried out on emotion recognition from facial expression with stimulus materials of isolated faces. However, faces seldom appear in isolation. They usually appear in a situational scene, with body language, and/or with voice. Since Ekman's theory predominates in the field of facial expressions, the role of contextual information on recognising emotion from facial expressions may be underestimated.

Recently, researchers (Adams, Ambady, Macrae, & Kleck, 2006; Adams & Kleck, 2003, 2005; Aviezer, Bentin, Dudarev, & hassin, 2011; Aviezer et al., 2008b; de Gelder, Böcker, Tuomainen, Hensen, & Vroomen, 1999; de Gelder & Vroomen, 2000; Dolan, Morris, & de Gelder, 2001; Meeren, van Heijnsbergen, & de Gelder, 2005; Righart & de Gelder, 2008b; Sander, Grandjean, Kaiser, Wehrle, & Scherer, 2007; van den Stock, Righart, & de Gelder, 2007) have turned their interests into placing faces back into contexts to find out whether contextual information plays a role in perceiving and recognising emotional facial expressions and to figure out to what extent contextual information influences emotional facial expression recognition.

The remainder of this chapter will discuss the universality of facial expressions of emotion, the theories of emotion, context effects in the recognition of facial expressions, and whether attentional resources are required for context effects. The last part of this chapter will concisely describe the main goals of this thesis.

1.1 The Universality of Facial Expression of emotions

It has been over a hundred years since emotions were studied. During this period, several psychological models of emotion have been proposed, which emphasise different aspects of emotions，such as emotional experience, cognitive antecedents of emotion, and physiological and expressive consequences of emotion. Among these theories, it is Basic Emotion Theory (Ekman, 1992) that is closely linked to the study of emotional facial expression.

The origin of Basic Emotion Theory comes from Charles Darwin who laid the foundation for research of facial expressions. Darwin (1872) suggested that both human and animals have emotional expressions that are considered as evolutionary product. Darwin conducted research - he sent research material and questions to Englishmen who lived abroad in eight different cultures - to support his assumption on universality of emotional facial expressions (as cited in Ekman, 1999). It was reported that these Englishmen could recognise the facial expressions shown by the natives and the facial expressions conveyed by the natives were the same to those by people in England. Thus, he suggested that these expressions are biologically determined and can be recognised universally, and cross-culturally. The method of asking participants to judge emotions from photographs is still used.

Darwin’s idea was extended by several other emotion theorists (Izard, 1971; Ekman, 1972; Oatley & Johnson-Laird, 1987; Plutchik, 1962, 1980; Tomkins, 1962). These theorists held the similar view that some emotions are basic and primary, while differences exist regarding how many and which emotions are basic. The number of basic emotions varies from 2 to 18. To date, the notion of six basic emotions (happiness, surprise, fear, anger, disgust, sadness) proposed by Ekman (1972) are the most influential and most accepted. These six basic emotions are closely associated with and are recognised from specific and characteristic facial expressions (Ekman, 1972; Ortony & Turner, 1990). The argument about whether these six emotional facial expressions are universal has been tested for a long time. Ekman and his colleagues, and other independent researchers conducted cross-cultural studies in 21 literate countries, including Western and non-Western countries, to test universality of emotional facial expressions (Boucher & Carlson, 1980; Ducci, Arcuri, Georgis, & Sineshaw, 1982; Ekman, Sorenson, & Friesen, 1969; Ekman et al., 1987; Izard,1971; Niit & Valsiner, 1977; McAndrew, 1986). In these studies, participants were asked to judge the emotion shown on each of the photographs. The options of emotions were on a list, translated in their own language. The results were in high agreement with Ekman's postulation, that is, these emotional facial expressions were recognised universally, and were not culturally variant.

However, one limitation to the studies of universality was that participants in these studies might have been exposed to Western mass media, so they might have already learned these emotional facial expressions. In other words, the conclusion Ekman and coworkers drew might be influenced by this factor. In order to rule out this possibility, another study was conducted in a preliterate, visually isolated cultures, the tribe of the South Fore, New Guinea (Ekman & Friesen, 1971). They set up the criteria to select appropriate participants. One group of participants selected did not have any chance to fully learn Western facial expressions, and the other group of participants selected were in frequent touch with Western culture, such as all spoke English, had seen Western movies, and worked for Western government. Because the group of participants who had little contact with the Westerners could not read, the procedure was changed: the stories were read to the observers, and then they needed to choose one that had the same emotional meaning as stories, from the three pictures. The analysis performed on the number of correct judgement and showed that there was on significant difference between these two groups. The results revealed that the experience of contact with Western culture did not make any difference of recognising the facial expressions between these two groups of participants. They could discriminate all basic emotions from each other, except that they could not discriminate fear from surprise. Heider and Rosch who carried out a study in a more visually isolated tribe of Dani, New Guinea replicated the results (as cited in Ekman et al., 1987). Heider and Rosch used the same procedure in their study, and arrived at the similar conclusion to that in Ekman and Friesen's study (1971). However, one distinction that was found in this study was that the Dani did not discriminate anger from disgust, while the South Fore did not discriminate fear from surprise. Together, the way in which basic emotions were interpreted for preliterate, visually isolated people was almost the same to that for literate people. People in preliterate, visually isolated culture had no obstacles to recognising emotional facial expressions of Western people. Then how about Western people? Do they understand the emotion facial expressions of people in preliterate culture? An experiment (as cited in Ekman &

Friesen, 1971) was designed to answer this question. If emotional facial expressions were universal, the answer to the question would be positive. In New Guinea, a group of people was asked to display their face when they were in each of the emotional situations the experimenter described. Observers in America were shown the videotapes of people from the visual-isolated culture, and were required to judge the emotion. The results showed that Americans who had never had contact with people in New Guinea could correctly judge their emotional facial expressions: anger, sadness, happiness and disgust, whereas fear and surprise were not distinguished from each other.

The empirical studies showed us that happiness, sadness, anger, disgust, fear, and surprise were recognised as being distinct from each other in the literate cultures, and the similar results were observed in the preliterate, visually isolated cultures. Thus, the collected evidence supported the notion that six basic facial expressions are recognised universally. However, Russell (1994) put forward the criticism of the method used in the studies mentioned above, in which participants were asked to choose from a list. This forced-choice format might have indirectly elicited the universality of facial expressions of emotion. In fact, he might ignore the existing studies in which participants were allowed to freely choose their own words, instead of using forced-choice task, to categorise the stimuli presented (Boucher & Carlson, 1980; Izard, 1971; Rosenberg & Ekman, 1994). The results of these studies revealed that participants chose the similar words to those provided in the studies that adopted the forced-choice task, suggesting that facial expressions are recognised universally regardless of the method used. Another powerful piece of evidence was provided by the comparison of spontaneous facial expressions between American and Japanese (as cited in Ekman, 1972). When these two groups of participants were watching films (stress-inducing films and neutral films), their facial expressions were recorded. When they were alone in the room, with a hidden camera, the facial expressions displayed were the same for both American and Japanese; while when they were with the experimenter, Japanese were more likely to hide their facial expressions. The differences of facial expressions between American and Japanese might be explained by "display rules" which were defined as "procedures learned early in life for the management of affect displays and include deintensifying, intensifying, neutralizing, or masking an affect display" (Ekman et al., 1969, p.69). In other words, in private situations, facial expressions shown on Americans and Japanese observers are with no difference. However, in public situations, Japanese tend to mask negative facial expressions, whereas Americans do not. In general, facial expressions is universal and is not culturally specific. Where there are cultural differences is how we control and manage our facial expressions.

1.2 The Theories of Emotion regarding Facial Expression Recognition

A long and debated question in the field of emotion is whether emotion is perceived as discrete categories or as continuity along one or more dimensions (Ekman, 1992;

Russell, 1997; Schlosberg, 1954). Discrete emotion theories proposed that there are a limited number of distinct, separate emotions which provide important signals for survival (Ortony & Turner, 1990). Basic emotion theory, one of the influential discrete emotion theories, at first suggested that there are six basic emotions, later adding contempt into the list. According to basic emotion theory, a specific pattern of facial muscular movement represents a certain category of emotion. These basic emotions are different from one to another, as each of the six basic emotions possesses a universal and unique facial expression (Ekman, 1999) and a specific pattern of autonomic nervous system (ANS) activity (Kreibig, 2010; Lench, Flores, & Bench, 2011). The structural information of a basic facial expression is sufficient for individuals to classify the facial expression into a certain emotional category, which would not be influenced by contextual information (Fugate, 2013). For example, the brows lowered and drawn together, the upper eyelid raised and the muscle in the lips tightened is categorised as anger, rather than any other emotion categories.

By contrast, Dimensional theories argued against the idea that the boundary between emotions is clear, instead stating that emotions can be located into a range that is either a unidimensional space or multidimensional space (Scherer, 2000). The circumplex model (one of dimensional theories) proposed that there are two kinds of information delivered by faces: (a) at first, an observer perceive quasi-physical information (such as, smile, tears, and muscular contracts, etc.); (b) the observer judges an emotion of a face on the two dimensions of valence and arousal, based partly on the quasi-physical information. According to the circumplex model of emotion, emotions are roughly placed on a circle with the two dimensions of valence and arousal (Russell, 1980, 2003). Valence is a horizontal axis, ranging from unpleasant (misery, unhappiness) at one end through a neutral point to pleasant (happiness, contentment) at the other end. And, arousal is a vertical axis, ranging from deactivation (drowsiness, lethargy) through the same neutral point to activation (agitation, frenzy) (see Figure 1.1). From the dimensional view, facial expressions do not convey specific emotional information, but only deliver information mentioned above.

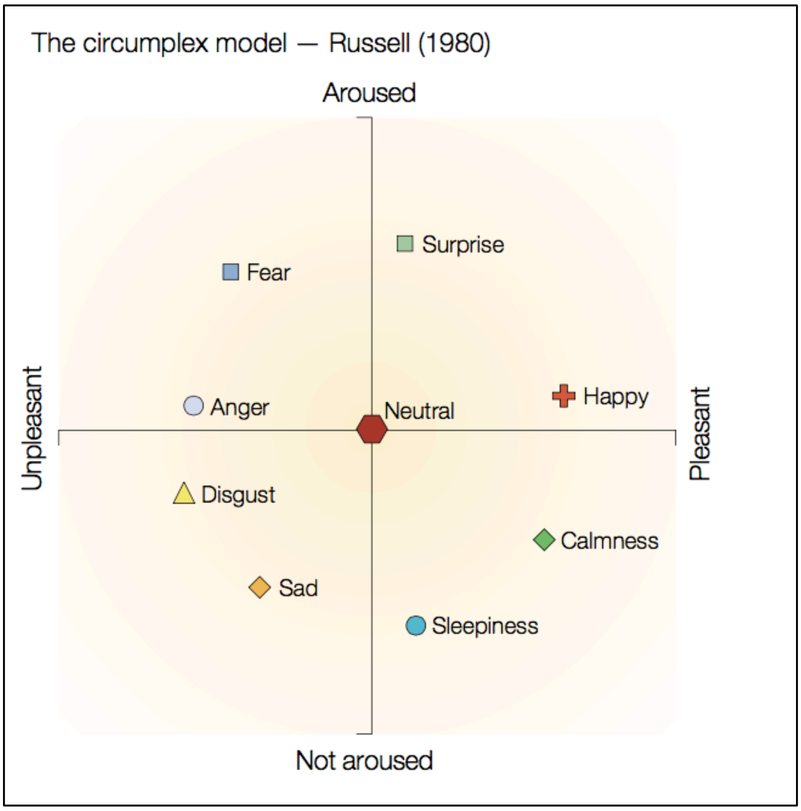

Figure 1.1. The circumplex model (1980): A schematic representation of the location of basic emotions basing on the two dimensions of arousal (vertical axis) and valence (horizontal axis). Reproduced from Calder, Lawrence, and Young (2001)

The notion of discrete categories of emotions (Basic Emotion Theory) suggests categorical perception of facial expressions (Ekman, 1994), whereas Dimensional Theories (the circumplex model) do not. A number of categorical judgment studies were carried out to test whether the perception of emotional facial expression is under the categorical way or the dimensional way. Studies on categorical perception have demonstrated that categorical perception occurs when humans classify a stimulus varying continuously along one or more dimensions into one of discrete, separate categories (Goldstone & Hendrickson, 2010; Harnad, 1987; Young et al., 1997). According to the view of categorical perception, it is easier to discriminate differences between a pair of stimuli that locate in two different categories (between-category discrimination) than between a pair that locates in the same category (within-category discrimination) (Harnad, 1987; Young et al., 1997).

Etcoff and Magee (1992) were the first to examine categorical perception of facial expressions of emotion. They used Brennan's (1985) algorithm to create a series of facial expressions morphing from one prototypical facial expression to another prototypical facial expression (e.g., from anger to sadness, or from anger to fear) or from one prototypical facial expression to a neutral face in equal physical stages. Each series included 11 stimuli. The stimuli of faces used in this study were line drawings of photos. It was hypothesized that if the perception of expressions was not categorical, the linear trend of the identification of facial expressions would be observed, whereas if the perception of expressions was categorical, the percentage of the identification of facial expression would change abruptly at a certain boundary. Participants were asked to complete the identification task and the discrimination task. In the identification task, participants were asked to press one of two keys to indicate the emotion shown on the face (e.g., anger or fear). The trend analyses (including a linear contrast, the deviation from the linear contrast, and a contrast orthogonal to the linear trend) were performed. The results showed that there was a clear boundary between facial expressions in each series, with participants classifying them into different categories. However, categorical perception of surprised facial expressions was not found. In the AB-X discrimination task, three faces, A, B, and X, were sequentially presented to participants, and the first two stimuli (A and B) were usually different from each other. Then participants were asked to press a key to indicate whether X was the same as A or B. Etcoff and Magee set a criterion for the boundary to divide facial expressions of each series into two categories. The results showed that discrimination of between-category faces was more accurate than that of within-category faces, which showed the typical pattern of categorical perception of facial expressions. As in the identification task, categorical perception was not found for surprised expressions in the discrimination task.

Other studies (Calder, Young, Perrett, Etcoff, & Rowland, 1996; de Gelder, Teunisse, & Benson, 1997; Young et al., 1997) used real photographs instead of line drawings, and used a more advanced morphing technique (Benson & Perrett, 1991) to create continuum-stimuli between two prototypical facial expressions. Calder et al. (1996)

carried out four experiments to replicate Etcoff and Magee's findings. The first two experiments used the identification and discrimination tasks, similar to the procedure of Etcoff and Magee (1992). In order to avoid range effects[1], Experiment 3 used a special continuum from three prototypical expressions (fear→happiness→anger→fear), and the three prototype stimuli were not showed to participants. Experiment 4 used the same stimuli as those in Experiment 3, and adopted the same-different matching paradigm in which the two facial expressions were presented to participants at the same time to address the disadvantage of short-term memory in the AB-X discrimination task. All these four experiments replicated Etcoff and Magge's findings, showing that the patterns of participants' data were consistent with categorical perception of facial expressions. The results showed that a boundary was observed in each continuum in the identification task, and between-category pairs were discriminated better than within-category pairs. Young et al. (1997) created all possible pairs of the six basic emotions. The patterns of identification rates and discrimination performance both suggested that the facial expressions were recognised categorically, which was reflected in a sudden shift in the middle of each continuum in the identification task and with better discrimination across boundaries and poorer within-category discrimination performance. Evidence of categorical perception was not only found for adults, but also for children, both of them showing better

---

[1] Range effects: a consequence of using single continua ranging between two prototypical facial expressions.

performance in cross-boundary pairs than in within-category pairs (de Gelder et al., 1997). The categorical perception effects were even found for 7-month-old infants (Kotsoni, De Haan, & Johnson, 2001).

However, the method adopted in the studies demonstrating categorical perception of facial expressions was challenged. It is argued that the categorical perception of facial expressions was only derived from the results of the identification and discrimination tasks. To rule out the possibility, Bimler and Kirkland (2001) used multidimensional scale (MDS), which was commonly adopted to demonstrate dimensional view of facial expressions, to examine categorical perception of facial expressions. MDS was used to produce spatial models and non-spatial models to explore the similarities among the perceived emotional faces. Bimler and Kirkland (2001) used blends of four basic facial expressions (angry, sad, surprised, and happy) and neutral faces to create series of morphed facial expressions. Participants were instructed to indicate proximities between these facial expressions rather than categorising facial expressions. The data were analysed using MDS and showed that adjacent facial expressions were clustered within each category. The results suggested facial expressions are perceived in discrete categories.

Contrary to categorical perception of facial expressions, Takehara and Suzuki (1997) asked participants to rate morphed facial expressions that were similar to those used in the study carried out by Etcoff and Magee (1992) on the 9-point rating scale (from "not at all" to "very much so"). They used multidimensional scale (MDS) to analyse the data. Their results revealed that facial expressions were recognised along valence and arousal dimensions, but were not recognised in cluster. The findings were replicated by other studies (Katsikitis, 1997; Takehara & Suzuki, 2001). Later on, several studies (Christie & Friedman, 2004; Fujimura, Matsuda, Katahira, Okada, & Okanoya, 2012; Panayiotou, 2008; Russell, 2003) have found that categorical and dimensional perception of facial expressions might co-occur. Fujimura et al. (2012) used a 9 x 9 affect grid to evaluate emotional faces along horizontal valence and vertical arousal dimensions. The 9 x 9 affect grid is another method, which is used to test the dimensional view of facial expressions. They observed that morphed facial expressions were clustered into two groups either based on valence or arousal, indicating that categorical perception effects occurred under the dimensional strategy. This finding suggested a hybrid theory that both categorical perception and dimensional perception of facial expressions were used simultaneously.

To summarise, it is inconclusive whether facial expressions are recognised in discrete categories or in valence and arousal dimensions. Regardless, Basic Emotion Theory has become one of the predominant models in the area of perception of emotional facial expressions.

1.3 Context Effects on Recognition of Facial Expressions

The prevalence of discrete emotion theories has led to isolated faces often being used in the domain of emotion research. However, outside the laboratory, faces are usually perceived with other social cues, including visual context cues (e.g., body language, situational scene, and gaze direction) and auditory context cues (e.g., verbal information depicting situational scene, tone of voice with neutral verbal information, and tone of voice without any verbal information), which may have an impact on how they are perceived. The following sections will review and discuss the visual and auditory context effects on recognising facial expressions.

1.3.1 Visual Context Effects and Facial Expression Recognition

1.3.1.1 Body Language

Recently, researchers have become interested in perception and recognition of emotion from body language (Atkinson, Dittrich, Gemmill, & Young, 2004; de Gelder, 2006; de Gelder & van den Stock, 2011; Grezes, Pichon, & de Gelder, 2007; Schindler, Gool, & de Gelder, 2008;). It was suggested that body language, as facial expressions, may provide reliable cues for emotion recognition, and that when facial expressions are simultaneously presented with bodies, the recognition of facial expressions would be influenced by the bodily expressions (Meeren et al., 2005; van de Riet & de Gelder, 2008; van den Stock et al., 2007).

Meeren et al. (2005) investigated the influence of emotional body language on the recognition of facial expressions. In this study, face-body compound (see Figure1.2), consisting of images of faces and whole bodies displaying fearful or angry, were either emotionally congruent or incongruent. Participants were asked to judge whether the face expressed fear or anger. It was found that participants judged facial expressions more accurately and faster when seeing the emotional-congruent compound than the emotional-incongruent pairs. Compared to congruent pairs, the incongruent bodies impaired the recognition of facial expressions. van den Stock et al. (2007) used fearful and happy facial expressions to test the influence of body language on the recognition of facial expressions. Fearful and happy faces were combined with fearful and happy bodies to create either congruent pairs or incongruent pairs. Fearful faces were categorised as happy more frequently when accompanied with happy bodies than with fearful bodies, which replicated the results of Meeren et al.'s study (2005). However, there were no significant differences of percentage of categorisation a happy facial expression as happy between a happy face paired with a happy body and a happy face paired with a fearful body, which indicated that the recognition of happy faces was not influenced by emotions shown by bodies. Fearful and happy faces were used and tested in another study (van de Riet & de Gelder, 2008). The face stimuli and the body stimuli were edited into the same height, and were superimposed on each other (see

Figure 1.3). The results showed that the accuracy of recognition of fearful faces was not affected by body expressions, whereas happy faces were recognised better when paired with happy bodies than when paired with fearful bodies. Contrary to Meeren et al.'s (2005) findings, reaction times for recognising facial expressions were not influenced by the bodily expressions.

These studies reviewed in this subsection have shown that the emotion conveyed by body language influenced the recognition of facial expressions, with significant differences found between facial expressions with congruent bodies and facial expressions with incongruent bodies. Although the recognition of facial expressions was influenced by concurrently presented body language, the patterns of context effects were not the same. For example, the better recognition for fearful congruent face-body stimuli was found in some studies (Meeren et al., 2005; van den Stock et al., 2007), but was not observed in the study of van de Reit and de Gelder (2008). The different patterns of context effects might be elicited by different types of face-body compound stimuli (as shown in the Figure 1.2 and 1.3). The size of an affective stimulus can influence the accuracy and reaction times of judgment task, with lower accuracies and longer response times for smaller, negative stimuli (De Cesarei & Codispoti, 2006). Thus, the relative smaller bodies in the study of van de Reit and de Gelder (2008) elicited different patterns of context effects on the recognition of facial expressions. In these cited studies, statistical analyses were not carried out between paired-stimuli and isolated face stimuli. Hence, it is unknown whether there are facilitation effects for the congruent condition compared to isolated faces or interference effects for the incongruent condition by contrast with isolated faces.

Figure1.2. Examples of compound stimuli adopted in the study of Meeren et al. (2005). The left part was congruent pairs, an anger face with an anger body, and a fear face with a fear body; the right part was incongruent pairs, an anger face with a fear body, and a fear face with an anger body. Reproduced from Meeren et al. (2005).

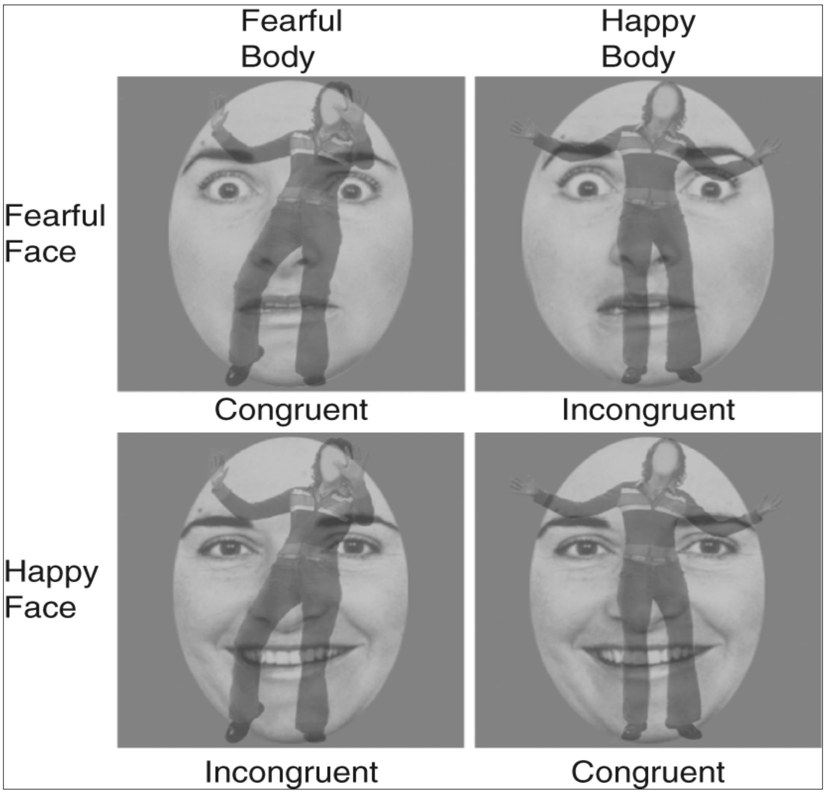

Figure 1.3. Examples of compound stimuli used in the study of van de Riet and de Gelder (2008). The top part presented a fearful face paired with a fearful congruent body (shown in top left) and with a happy incongruent body (in top right). The bottom part presented a happy face paired with a fearful incongruent body (in bottom left) and a happy congruent body (in bottom right). Reproduced from van de Riet and de Gelder (2008).

1.3.1.2 Visual Scenes

In addition to body language, the accompanied emotional visual scenes also have an impact on the recognition of facial expressions (Lee, Choi, & Cho, 2012; Righart & de Gelder, 2008a, 2008b).

Righart and de Gelder (2008b) attempted to demonstrate the influence of emotional scenes on the recognition of facial expressions. In Experiment 1A, faces and scenes were combined to create compound stimuli, with 3 congruent pairs (fear-fear (FF), disgust-disgust (DD), and happy-happy (HH)) and 6 incongruent pairs (FD, FH; DF, DH; HD, HF). Participants were asked to complete a three-alternative forced choice task, indicating whether the emotion expressed by the face was fear, anger, or happiness. In the Experiment 1B, another group of participants were instructed to indicate the facial expression in a two-alternative forced choice task in three independent blocks (Fear-Disgust block, Fear-Happiness block, Disgust-Happiness block). The results from experiment 1A and 1B showed that the congruency effects were found for happy facial expressions, with fewer errors and faster RTs for happy faces in happy scenes than for happy faces in fearful scenes, and for happy faces in disgusted scenes. The recognition of fearful facial expressions was better when paired with congruent fearful scenes than that of fearful facial expressions when paired with incongruent disgusted scenes, but was not better than that of fearful facial expressions when paired with incongruent happy scenes. The results suggested that the recognition of fearful and happy facial expressions is influenced by accompanied emotional scenes. In terms of disgusted facial expressions, experiment 1A obtained fewer error rates and faster reaction times for the recognition of disgusted faces in emotionally congruent scenes compared to those for the recognition of disgusted faces in emotionally incongruent scenes, whereas experiment 1B did not. It was suggested that the results might result from the number of response options provided. Carroll and Russell (1996) stated that participants made a response more on the simultaneously presented context information when more alternatives were provided.

In another study, Righart and de Gelder (2008a) examined how the early stages of facial expression processing are influenced by emotional scenes. A facial expression (fear, happiness) was paired with each category of emotional scenes (fear, happiness, neutrality). Participants performed a two-alternative forced choice task in which they judged whether the face was happy or fearful. The behavioural data showed that reaction times for the recognition of a happy facial expression were faster in a happy and a neutral scene than those for the recognition of a happy facial expression in a fearful scene, but no significant differences of reaction times were found between the recognition of a happy face in a happy scene and that of a happy face in a neutral scene. The results of the happy faces replicated the study of Righart and de Gelder (2008b), suggesting incongruent scenes impairs the recognition of happy faces. However, contrary to the study by Righart and de Gelder (2008b), the recognition of fearful facial expressions was not influenced by accompanied emotional scenes, as there were no significant differences of reaction times for the recognition of fearful facial expressions when paired with fearful scenes, happy scenes, and neutral scenes.

Similar to emotional bodily expressions, it was found that the recognition of facial expressions was influenced by emotional scenes, though different patterns of context effects on facial expression recognition was obtained.

1.3.2 Auditory Context Effects and The Facial Expressions of Emotion

Humans can perceive emotion not only from facial expressions but also from voices (Ethofer, Pourtois, & Wildgruber, 2006a). Researchers have used behavioural, neuroanatomical, electrophysiological, and neuroimaging methods to investigate audiovisual integration of emotional signals and explore the effects of auditory information in the recognition of facial expressions.

de Gelder and Vroomen (2000) explored the combination of information from facial expressions and tones of voices in the recognition of emotion. In one of their experiments, happy and sad facial expressions and voices were adopted. Participants were presented with face-voice compounds (either emotionally congruent or incongruent), and were asked to judge whether the face was happy or sad, ignoring the tones of voices. The results showed that reaction times were slower in the incongruent compounds than in the congruent compounds, suggesting the incongruent voices impair the recognition of facial expressions. The results were replicated by another study (Dolan et al., 2001). In this study, Dolan et al. (2001) used event-related functional Magnetic Resonance Imaging (fMRI) to examine how the presentation of voices influences the recognition of facial expressions. The experiments used a computer-morphing procedure to create fearful facial expressions of level of 50% physical intensity and happy facial expressions of level of 50% physical intensity. The face-voice compounds consisted of the morphed emotional faces (fearful or happy) and neutral sentences with a fearful tone or a happy tone. Participants were asked to categorise the facial expressions as either fearful or happy as quickly and accurately as possible. Behavioural data indicated that reaction times of the congruent face-voice stimuli were faster than those of the incongruent ones. Moreover, the greater activation of left amygdala was found for the congruent face-voice stimuli than for the incongruent ones. Further analyses showed that the activation of amygdala was mainly elicited by fearful congruent pairs. Besides the amygdala, the fearful face-fearful tone stimuli elicited an activation of right fusiform cortex, compared to fearful face-happy tone.

The results were replicated by a study performed by Ethofer et al. (2006b), in which participants were asked to rate the valence of emotional faces on a nine-point self-assessment manikin scale (SAM), while ignoring concurrently presented voices. The blood oxygenation level-dependent (BOLD) responses in the right fusiform gyrus were enhanced when participants rated fearful faces simultaneously presented with congruent fearful tones, in contrast with when participants rated the same fearful faces with incongruent happy tones. Another fMRI study (Kreifelts, Ethofer, Grodd, Erb, & Wildgruber, 2007) investigating on the audiovisual integration of emotional signals in voices and faces revealed that the bilateral pSTG and right thalamus were more strongly activated by the audiovisual stimuli than the unimodal stimuli.

In a recent neuroimaging study, Müller et al. (2011) explored the influence of emotional sounds on facial expressions perception and neural structures using audiovisual congruent and incongruent stimuli. Due to the fact that no context effects were found for prototypical fearful and happy facial expressions in their pre-study, the degraded happy or fearful faces were created by merging the prototypical fearful or happy facial expressions with the neutral mouths of the same actors. The face-voice stimuli were comprised of faces (either displaying happy, neutral, or fearful) and auditory stimuli. The auditory stimuli were either yawn (neutral) or emotional sounds like laugh (happy) or scream (fearful). Participants were instructed to ignore the sounds and to rate the presented facial expressions on an eight-point rating scale from extremely fearful to extremely happy. The results revealed that degraded fearful facial expressions were rated as being more fearful when accompanied by screams compared to yawns or laughs. There were no context effects found for degraded happy facial expressions. These results were consistent with the results of Ethofer et al. (2006b) who also found context effects for fearful but not for happy faces. Happy facial expressions were recognised more accurately and faster than any other facial expressions (Ekman, Friesen, & Ellsworth, 1982; Gur et al., 2002; Kirita & Endo, 1995; Montagne, Kessels, de Haan, & Perrett, 2007). This happy face advantage might be the reason that there were no context effects on the recognition of happy facial expressions. Moreover, the imaging data showed that incongruence between emotional faces and sounds evoked greater activation in the middle cingulate cortex, right superior frontal cortex, right supplementary motor area as well as the right temporoparietal junction. Taken together, the empirical data from these mentioned audiovisual studies revealed better recognition performance and larger activations for congruent face-sound stimuli, suggesting that concurrently presented voice plays an important role in the recognition of facial expressions.

In sum, the above-mentioned studies have found that emotional context information has an impact on the recognition of facial expressions. However, there is no consistent pattern of the contextual effects. For example, de Gelder and Vroomen (2000) found that happy facial expressions were recognised more accurately and faster for the congruent pairs than for the incongruent pairs, whereas these effects were not found in other studies (Ethofer et al., 2006b; Müller et al., 2011). The different patterns were observed for the context effects on the recognition of other facial expressions. The inconsistent pattern of context effects might be explained by emotion seed view (Aviezer et al., 2008a, 2008b), which will be discussed in detail in the following section.

1.3.3 Emotional Seeds View

Aviezer et al. (2008a, 2008b) put forward the emotion seeds view to predict when context effects would occur and the magnitude of contextual effects. Before elaborating this view, I will introduce the term of "emotion seeds".

Each facial expression has its own unique configuration that consists of physical features (e.g., face muscular movements). Some of these physical features are special for only one category of emotional facial expression, such as, the wrinkled nose in disgusted faces, while other features are shared with one or more other expressions. For instance, the furrowed brows are common in disgusted and angry faces, and the raised brows and wide eyes are shared with fear and surprise faces (Ekman & Friesen, 1976, 1978). These shared physical features are called emotion seeds. The more emotion seeds the two facial expressions share, the greater these two facial expressions resemble each other. Aviezer et al. (2008a, 2008b) argued that when a face appears in isolation, these shared physical characteristics (emotion seeds) remain inactive and have little impact on the recognition of facial expressions, whereas when a face is presented in a context that conveys an emotion similar to that delivered by the face, these seeds would be activated to influence the recognition of the facial expressions. Emotion seeds in faces refer to perceptual similarity between emotional facial expressions and perceptual similarity can be evaluated and calculated via multidimensional scaling (MDS) (Dailey, Cottrell, Padgett, & Adolphs, 2002; Susskind, Littlewort, Bartlett, Movellan, & Anderson, 2007). Figure 1.4 illustrates perceptual similarity between six basic facial expressions. Aviezer et al. (2008a, 2008b) proposed that the largest magnitude of context effects happens when an emotion displayed by a facial expression is highly similar to an emotion conveyed by a context with which the face appears. For example, disgusted facial expressions are perceived to be the most similar to angry faces, so angry contexts evoke the greatest interference effects on the perception of disgusted faces. Sadness, which is moderately similar to disgust, evoke moderate interference effects on the perception of disgusted faces. And fear, which is the least similar to disgust, evoke the smallest interference effects on the perception of disgusted facial expressions.

Figure 1.4. The position of six basic emotional facial expressions. The adjacent emotion categories are more physically similar to each other. Reprinted from Aviezer et al. (2008a)

Aviezer et al. (2008b) performed a study that lent support to the emotion seeds view. Disgusted facial expressions were placed on images of emotional contexts (upper bodies) expressing disgusted, angry, sad, or fearful emotion. Besides body language, some of body images included props (e.g., a dirty diaper in a disgusted body image, and a gravestone in a sad body image). These emotional facial expressions and bodily expressions were combined together to create four levels of perceptual similarity: (a) disgust-disgust (full similarity), (b) disgust-anger (high similarity), (c) disgust-sadness (medium similarity), and (d) disgust-fear (low similarity). Participants were presented with the face-body stimuli and a list of six basic-emotion labels (anger, disgust, fear, happiness, sadness, and surprise) under each of the stimuli. Participants were instructed to press a button to categorise the facial expression, ignoring the emotion shown by body language. The results showed that the mean accuracy of disgust faces was 91% when paired with disgusted bodies (full similarity), 59% when paired with fearful bodies (low similarity), 35% when paired with sad bodies (medium similarity), and 11% when paired with angry context (high similarity). That is, categorisation of disgusted faces was impaired the most when disgusted faces were paired with bodies displaying anger; was impaired the least when paired with bodies displaying fear; and was moderately impaired when paired with bodies displaying sadness. The findings were consistent with emotion seeds view which predicted that the magnitude of context effects varies with perceptual similarity between emotions conveyed by a face and by a context.

Some other studies (Aviezer, Trope, & Todorov, 2012; Mondloch, Nelson, & Horner, 2013b) were carried out to test whether context effects on recognition of facial expressions are consistent with emotion seeds view. In the study of Aviezer et al. (2012), they used four prototypical facial expressions (anger, disgust, sadness, and fear) instead of only one facial expression (disgust) in the previous study (Aviezer et al., 2008b). Emotional contexts, which included upper bodies and emotional paraphernalia (e.g., a dirty diaper), conveyed the same four emotions as the faces. Participants were presented with face-body stimuli and were asked to judge the expression on faces (anger, sadness, disgust, or fear), ignoring the emotion conveyed by bodily expressions. Of the four facial expressions, the recognition of disgusted and angry facial expressions was boosted by the congruent bodily expressions compared to isolated disgusted and angry facial expressions. However, these facilitation effects were not observed for the other two emotional faces (fear and sadness). The recognition of fearful and sad facial expressions was not improved by the congruent bodily expressions. In terms of interference effects, highly similar bodily expressions impaired the recognition of angry, disgusted, and sad facial expressions. In contrast, highly similar bodily expressions did not impair the recognition of fearful faces. The findings for fearful faces were not in line with the prediction of emotion seeds view which suggested the largest interference effects occur when a face and a context are highly similar.

Another attempt (Mondloch et al., 2013b) investigated whether the influence of body language on recognition of facial expressions is consistent with emotion seeds view. Emotional faces and body stimuli conveyed three emotions: anger, sadness, and fear. The body images adopted in this study were whole bodies, without any props. Each of the three facial expressions was paired with each of the three body expressions. Participants were asked to press a key to indicate what emotion depicted by the faces as quickly and accurately as possible. Of the three facial expressions, the pattern of context effects on the recognition of fearful facial expressions was consistent with emotion seeds view which predicted that greater interference effects would be elicited by sad bodies than angry bodies. The pattern of context effects on the recognition of the other two facial expressions (sadness and anger) was not in line with the prediction of emotion seeds view. According to this model, for sad facial expressions, angry bodies and fearful ones should lead to the equal magnitude of influence on the recognition of sad facial expressions. On the contrary, the results of this study showed that interference effects on recognition of sad facial expressions were observed when paired with fearful bodies, but were not found when paired with angry bodies. As for recognition of angry facial expressions, emotion seeds view predicted that contextual influence would be greater if angry faces were paired with sad bodies than paired with fearful bodies. However, the results of this study showed that fearful bodies produced interference effects, whereas sad bodies did not. The findings revealed that the pattern of context effects could be partially explained by emotion seeds view.

1.3.4 Summary

In short, the studies cited indicated that the recognition of some basic facial expressions benefits from the consistent contextual information (higher accuracy rates and faster reaction times), and is impaired by inconsistent contextual information (higher error rates and longer reaction times), while the recognition of other facial expressions is not influenced by contextual information. Emotion seeds view argued that it is the level of perceptual similarity between emotions conveyed by faces and contexts that elicit the inconsistent patterns of contextual effects on recognition of facial expressions. As predicted by emotion seeds view, the higher the perceptual similarity between facial expressions is, the greater the contextual effects on recognition of facial expressions are.

1.4 Ambiguous Facial Expressions and Context Effects

Two types of mixed facial expressions were created in the field of emotion research, which are referred as ambiguous facial expressions: (a) Type A, one prototypical facial expression is morphed into another prototypical facial expression at a certain step (e.g., 10% or 20%), thus creating a series of mixtures of two facial expressions (e.g., a series of mixed facial expressions changing from a 100% fearful expression to a 100% angry expression); (b) Type B, a series of mixed facial expressions is created by morphing between neutral faces and prototypical facial expressions, therefore producing different levels of intensity of facial expressions (e.g., a series of mixed facial expressions changing from a neutral expression to a 100% fearful expression).The ambiguous facial expressions were created by morphing software, with changing physical distances between key points of the features of facial expressions. For example, a level of 40% intensity of happy facial expression was produced by reducing the distance between the corners of the mouth and the corners of the eye to 40% of the maximum distance. The lower intensity of a facial expression was called subtle facial expression.

It was predicted that the magnitude of context effects depends not only on the similarity between the target facial expressions and the emotions elicited by the contexts (Aviezer et al., 2008a, 2008b), but also on the levels of ambiguity of facial expressions (Trope, 1986; van den Stock et al., 2007). Trope (1986) posited that context effects should be larger for ambiguous facial expressions than for unambiguous facial expressions (i.e. prototypical facial expressions). The ambiguous facial expressions mentioned in this prediction refer to the situation of Type A.

1.4.1 Context Effects and Type A Ambiguous Facial Expressions

The prediction that context effects depend on the level of ambiguity of facial expressions has been tested in two studies (van den Stock et al., 2007; de Gelder & Vroomen, 2000). One study (de Gelder & Vroomen, 2000) discussed this topic using the auditory context (affective prosody); the other one (van den Stock et al., 2007) used the visual context (body images). de Gelder and Vroomen (2000) adopted happy and sad facial expressions to produce a 11-step series of mixture expressions. Each step of the series was paired with a happy or sad auditory stimulus. Their results were consistent with the prediction, showing that context effects of voice on facial expression recognition were larger for the ambiguous mixed facial expressions than for the end of the facial expressions (100% sadness and 100% happiness). The findings were replicated in the study of van den Stock et al. (2007). They created a 5-step series of mixtures of fearful and happy facial expressions. The morphed faces were pasted on fearful or happy bodily expressions. The results indicated that context effects of the bodily expressions varied as a function of the levels of ambiguous facial expressions. The bodily expressions influenced the recognition of facial expressions the most when the face showed the most ambiguous emotion (e.g., 50% happiness-50% sadness). These two studies together have clearly demonstrated the prediction that the magnitude of context effects depends on the levels of ambiguous facial expressions, with greater influence on the most ambiguous faces.

1.4.2 Context Effects and Type B Ambiguous Facial Expressions

People, in real life, use both unambiguous and subtle facial expressions to communicate their emotions in human-to-human interaction and in human-to-computer interaction. Researchers and designers turned their interest into subtle expressivity, which is the major issue in the areas of interactive technology (Liu & Picard, 2003; Suzuki & Bartneck, 2003). It has been found that it is possible to communicate emotions using lower levels of intensity of facial expressions (Bartneck & Reichenbach, 2005; Hess, Blairy, & Kleck, 1997), though it is a hard task. Previous studies (de Gelder & Vroomen, 2000; van den Stock et al., 2007) have found that the magnitude of context effects on recognition of Type A ambiguous facial expressions depends on the levels of ambiguity of facial expressions, with greater context effects occurring for more ambiguous facial expressions. The most ambiguous emotional faces of Type A were the 50%-50% level of mixture of two facial expressions (e.g., 50% anger-50% fear), while the most ambiguous emotional faces of Type B were the lowest intensity of emotional faces adopted in a certain study. However, it is unclear whether the pattern of magnitude of context effects on recognition of Type A ambiguous facial expressions is applicable to the magnitude of context effects on recognition of Type B ambiguous facial expressions. In other words, it is unknown whether the magnitude of context effects relies on the levels of Type B ambiguous facial expressions, with the lower the intensity of facial expressions is, the larger the context effects will be observed.

In the following section, I will review the studies relating to the context effects and different levels of intensity of emotional facial expressions. As mentioned above, Dolan et al. (2001) presented a level of 50% intensity of fearful or happy facial expression with either congruent or incongruent affective prosody. They found that reaction times for judging the facial expressions paired with congruent voices were faster than those paired with incongruent voices. The results revealed that emotional voices have an impact on the recognition of subtle facial expressions, which was replicated by Lee et al. (2012).

Lee et al. (2012) examined how contextual stimuli influenced the recognition of different levels of intensity of fearful facial expressions. In this study, they produced the face images morphing from neutral (0%) to fearful (100%) using a 10%-increment step. The face images were placed on the centre of the negative, positive, or neutral background pictures, which were selected from the International Affective Picture

System (IAPS) (Lang, Bradley, & Cuthbert, 1999). The face-background stimuli were presented to participants and they were asked to judge whether the faces were fearful or neutral. The percentage of accurate recognition of fearful faces was higher when paired with negative background images than when paired with neutral ones. And the percentage of accurate recognition of fearful faces was lower when paired with positive background images than when paired with neutral ones. The results showed that context effects occurred for fearful facial expressions on the levels of intensity from 30% to 60%, whereas context effects were not found for fearful facial expressions on the levels of intensity from 0 to 20% and from 70% to 100%. However, this study did not compare the magnitude of context effects across the levels of intensity from 30% to 60%, so it is unclear whether the magnitude of context effects varies with the levels of intensity.

In addition to using a recognition task, context effects on recognition of subtle facial expressions were also explored adopting a rating task. The study (Müller et al., 2011), above mentioned, asked participants to rate degraded facial expressions of fear and happiness on an eight-point rating scale from extremely fearful to extremely happy. Degraded fearful facial expressions were rated more fearful when presented with screams (congruent pairs) compared to the same degraded fearful facial expressions when presented with laughs (incongruent pairs). However, such context effects were not found for degraded happy faces, which was not consistent with the results of Dolan et al.'s study (2001) where reaction times for level of 50% intensity of happy faces with happy prosodies were faster than those for level of 50% intensity of happy faces with fearful prosodies. Ethofer et al. (2006b) also indicated that context stimuli modulated the valence rating of fearful faces on each level of intensity used (25%, 50%, 75%, and 100%). They (Ethofer et al., 2006b) found that fearful facial expressions were rated more fearful when accompanied by fearfully spoken sentences compared with isolated fearful facial expressions. Conversely, context effects were not observed on any level of intensity of happy faces, which was consistent with the results of Müller et al. (2011).

The studies mentioned in this section revealed that emotional contextual stimuli modulated the recognition of subtle facial expressions, which was reflected by that (1) the accuracy of subtle facial expressions when paired with incongruent context was lower as compared to that of the same facial expressions when paired with congruent context, and (2) a fearful face was rated more fearful when it was paired with a congruent sound than the same isolated fearful face and the fearful face paired with an incongruent sound. As far as I know, only one study by Mondloch (2012) made comparisons between the magnitude of context effects on recognition of subtle facial expressions (40% intensity of sad faces and 40% intensity of fearful faces) and the magnitude of context effects on recognition of unambiguous facial expressions (100% intensity of sad faces and 100% intensity of fearful faces), though it was not the main purpose of this study. The results showed that there were no significant differences between the magnitude of context effects on recognition of 40% intensity facial expressions and the magnitude of context effects on recognition of 100% intensity facial expressions, which was inconsistent with the prediction that larger context effects would be found for more ambiguous facial expressions.

1.4.3 Summary

The reviewed studies provided empirical evidence that contextual stimuli do influence the recognition of ambiguous facial expressions, no matter which type of ambiguous facial expressions. The studies exploring the magnitude of context effects on Type A ambiguous facial expressions showed that greater context effects occur when facial expressions are more ambiguous.

However, the empirical data from the study of Type B ambiguous facial expressions (Mondloch, 2012) did not support the prediction that the magnitude of context effects depends on the levels of intensity of facial expressions. As mentioned, since it was not the main purpose of the study (Mondloch, 2012), Mondloch (2012) did not analyse the magnitude of context effects separately for fearful and sad facial expressions. Mondloch (2012) used two levels of intensity of facial expressions (40% intensity and 100% intensity) to explore the relationship between context effects and the levels of intensity of facial expressions. In view of the current situation, more studies should be carried out to test the prediction whether the magnitude of context effects relies on the levels of intensity of facial expressions, that is, larger context effects occur on lower levels of intensity of facial expressions.

1.5 Context Effects and Attentional Resources

A large body of research has contributed to exploring the question of whether the processing of emotional visual stimuli, including emotional facial expressions, is independent of attentional resources. It is an unresolved and controversial topic. While some researchers (Bradley, Cuthbert, & Lang, 1996; Buodo, Sarlo, & Palomba, 2002; Pereira et al., 2006; Schupp, Junghöfer, Weike, & Hamm, 2003) support the idea that emotional visual stimuli are processed without the requirement of attentional resources, and the processing of emotional stimuli occurs quickly and unconsciously, via a subcortical pathway (Vuilleumier, Armony, Driver, & Dolan, 2001), others (Pessoa & Adolphs, 2010; Pessoa & Adolphs, 2011) challenged this notion and have suggested that the processing of affective visual stimuli is not faster than that of other visual information and is dependent on attentional resources. Moreover, the findings in behavioural, fMRI and EEG studies (De Cesarei, Codispoti, & Schupp,2009; Erthal et al., 2005; Lim & Pessoa, 2008; Mitchell, Sundberg, & Reynolds,2007; Pessoa, McKenna, Gutierrez, & Ungerleider, 2002a; Pessoa, Padmala, & Morland, 2005; Schupp et al., 2007; Silver, Ress, & Heeger, 2007) showed that when attentional resources were occupied by other challenging tasks, the processing of emotional visual information was decreased or eliminated, such as that greater activation of the amygdala and visual cortex elicited by fearful faces than by neutral faces was eliminated when all attentional resources were occupied by another task (Pessoa et al., 2002a). These studies together demonstrated that the attentional resources might play a role in the processing of emotional visual information.

Another interesting and controversial issue is whether attentional resources are required for the processing of context effects on recognition of facial expressions, that is, whether the magnitude of context effects is modulated by attentional resources. Before discussing it, I will first elaborate load theory (Lavie & Tsal, 1994; Lavie, 1995).

1.5.1 Load Theory

We are exposed to a huge amount of information in daily life. Which information is processed depends on what we attend to. Selective attention is the mechanism that can select some stimuli to be processed while simultaneously ignoring other concurrent incoming stimuli. The crucial and longstanding unresolved issues in the area of attention are whether attentional selection acts as structural limitations (filtering mechanisms) or processing limitations (capacity/resources) and whether selective attention occurs early or late in processing.

Early selection theory (Broadbent, 1958) proposed that perception has a limited pool of resources. Attention acts as a filter, blocking irrelevant stimuli out, at an early stage of processing. Later on, Treisman (1960, 1964) proposed a revision to Broadbent's filter theory, arguing that the filter operates to attenuate irrelevant stimuli, instead of blocking them. In contrast to early selection theory, Deutsch and Deutsch (1963) assumed that perception has an unlimited pool of resources, thus all information, including relevant and irrelevant stimuli, is attended to. Selection takes place at a later stage of processing, after each stimulus is fully identified (Duncan, 1980). Early and late selection theories have assumed attention as a structural bottleneck in the information processing, and the dispute between early and late selection theories concerns which stage this structural bottleneck is located at, that is, when selective attention occurs. The capacity theory of attention assumed attention to be a general, flexible resources, rather than a structural filter (Kahneman, 1973). It was posited that there is a limited amount of attention and the capacity can be freely allocated to concurrent tasks (Kahneman, 1973). Different stages in information processing differ in the amount of attentional resources they require. It is suggested that early stages of information processing (e.g., sensory analysis) do not need attentional resources, whereas late stages of information processing require greater attentional resources (Kahneman, 1973). Load theory (Lavie, 1995; Lavie & Tsal, 1994), which applies the capacity theory of attention, was put forward to offer a resolution to the debate between early-selection mechanism and late-selection mechanism. Lavie and Tsal (1994) reviewed previous studies on attention and pointed out that the results in some studies supporting late-selection attention mechanism had been obtained under low perceptual load conditions, whereas the results in other studies supporting early-selection attention mechanism had been obtained under high perceptual load conditions. The load theory integrates the early-selection model and the late-selection model into one hybrid model and suggests that perceptual load is a major factor in determining which one of the mechanisms will be activated at a certain circumstance. The term *perceptual load* refers to either the number of units in a display or the nature of processing required for each unit. Load theory (Lavie, 1995; Lavie & Tsal, 1994) proposed two types of selective attention mechanisms, a passive perceptual selection mechanism and an active cognitive mechanism. According to load theory, these two types of mechanisms for selective attention have different impact on distractor effects from irrelevant distractors. It is assumed that early selection occurs when the perceptual load is high, while late selection occurs when perceptual load is low. Under high perceptual load, the demands of task-relevant stimuli would exhaust all perceptual capacity. No enough capacity is available for processing of task-irrelevant distractors, thus preventing distractor effects triggered by task-irrelevant distractors (the early selection mechanism). In contrast, in situations of low perceptual load, task-relevant stimuli do not take up all of the available perceptual capacity, and the remaining capacity would involuntarily 'spill over' to the perception of task-irrelevant distractors (the late selection mechanism). After relevant and irrelevant stimuli are perceived, the active cognitive mechanism takes effect, which has the opposite effects on distractor effects triggered by task-irrelevant distractors. If the active mechanisms, such as working memory, are overloaded, the ability to maintain the priorities of stimulus processing becomes weakened, and distractor effects by task-irrelevant distractors increases rather than decreases. Taken together, the perceptual mechanism predicts that distractor effects caused by task-irrelevant distractors should be larger when the level of perceptual load in processing task-relevant stimuli is low than when that is high, whereas the cognitive mechanism hypothesis predicts that distractor effects by task-irrelevant distractors increase under high working memory load, compared to low working memory load.

1.5.1.1 Effects of Perceptual Load on Irrelevant Distractor Perception

Figure 1.5 and Figure 1.6 provide examples of the manipulation of perceptual load used in the study of Lavie (1995). One method for the manipulation of perceptual load was the alteration of the number of units, as illustrated in Figure 1.5. The task was to press one of two designated keys to indicate whether the target was ***z*** or ***x***, ignoring the peripheral irrelevant letter. In the Condition S1, there was only one target letter and one irrelevant distractor letter presented to participants (the low perceptual load). In the Condition S6, a target letter with five other non-target letters and one irrelevant distractor letter were presented to participants (the high perceptual load). The irrelevant distractor was either compatible, or incompatible, or neutral with the target letter. The results of the experiment supported load theory, with greater distractor effects by irrelevant distractors being found under the low load condition, and with no distractor effects under the high load condition.

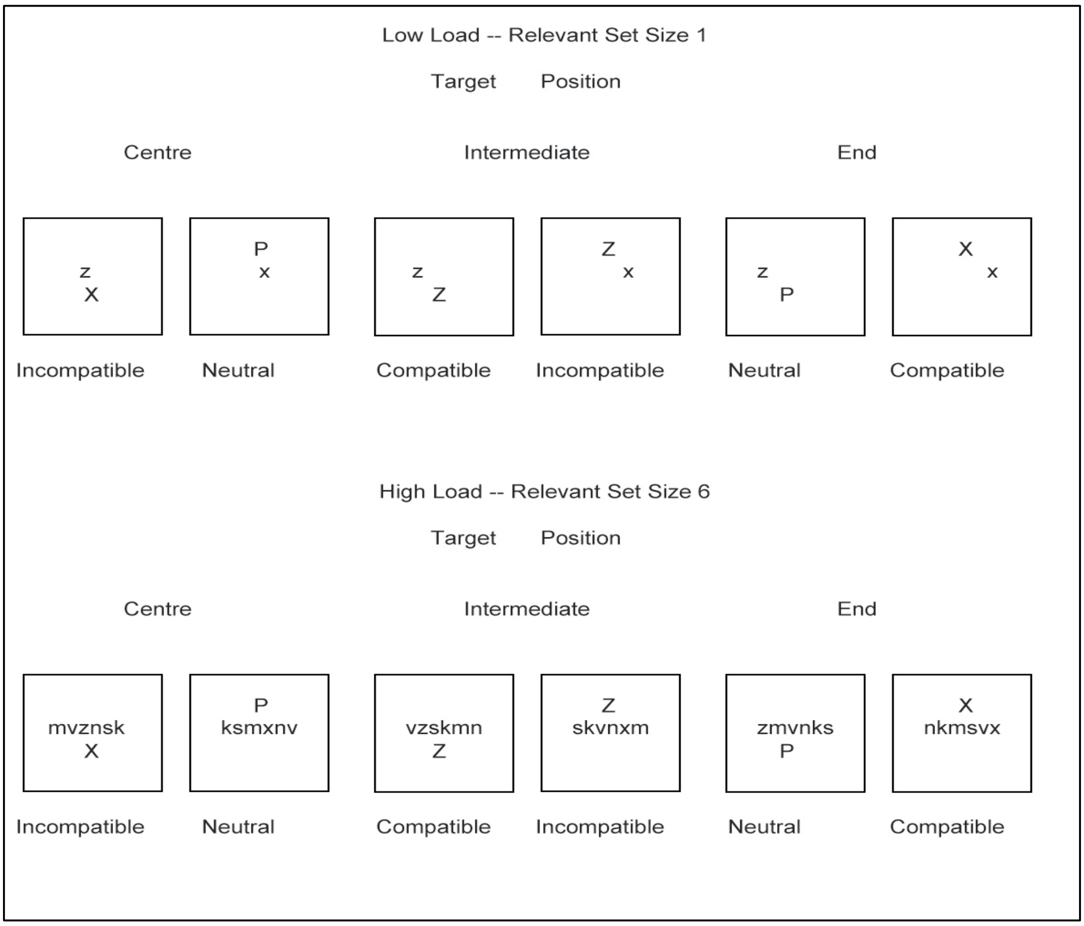

Figure 1.5. Examples of manipulation of perceptual load by altering the number of units. Adapted from Lavie (1995).

The perceptual load hypothesis was also demonstrated using another method of manipulating the levels of perceptual load, by increasing perceptual processing requirements for the same displays. The method to manipulate perceptual load was adapted from feature integration theory (Treisman & Gelade, 1980; Treisman & Sato, 1990), as illustrated in Figure 1.6. According to feature integration theory, individual feature perception (such as, shape, colour, and orientation) is independent of attentional resources, while combining individual features requires focused attention and therefore the level of perceptual load is increased. The task in Figure 1.6 used the go/no-go paradigm to test the hypothesis of load theory. Participants were asked to indicate whether the target letter was *H* or *U*, meanwhile ignoring the peripheral irrelevant letter. In the *feature demand condition* (low perceptual load), participants were instructed to make a response to the target letter when the colour of the additional shape was red (*Go* trials) and to withhold a response when it was blue (*No-Go trials*), irrespective of whether the shape was a circle or a square. In the *conjunction demand condition* (the high perceptual load), participants had to respond to the target letter when they saw a blue square or a red circle (*Go* trials), not to respond when they saw the opposite conjunction, a red square or a blue circle (*No-Go* trials). The distractor effects were merely observed under the low perceptual load condition, and not found under the high perceptual load condition, which is in line with the prediction of load theory. Generally, the degree of distractor effects from task-irrelevant distractors depends on the level of perceptual load imposed by the processing of relevant information. That is, when the demand of a perceptual relevant task is high, the perceptual capacity will be fully occupied by the processing of the relevant target, and then no sufficient capacity will be left for the irrelevant distractors. Given that irrelevant distractors are prevented from being processed, the distractor effects are decreased or eliminated.

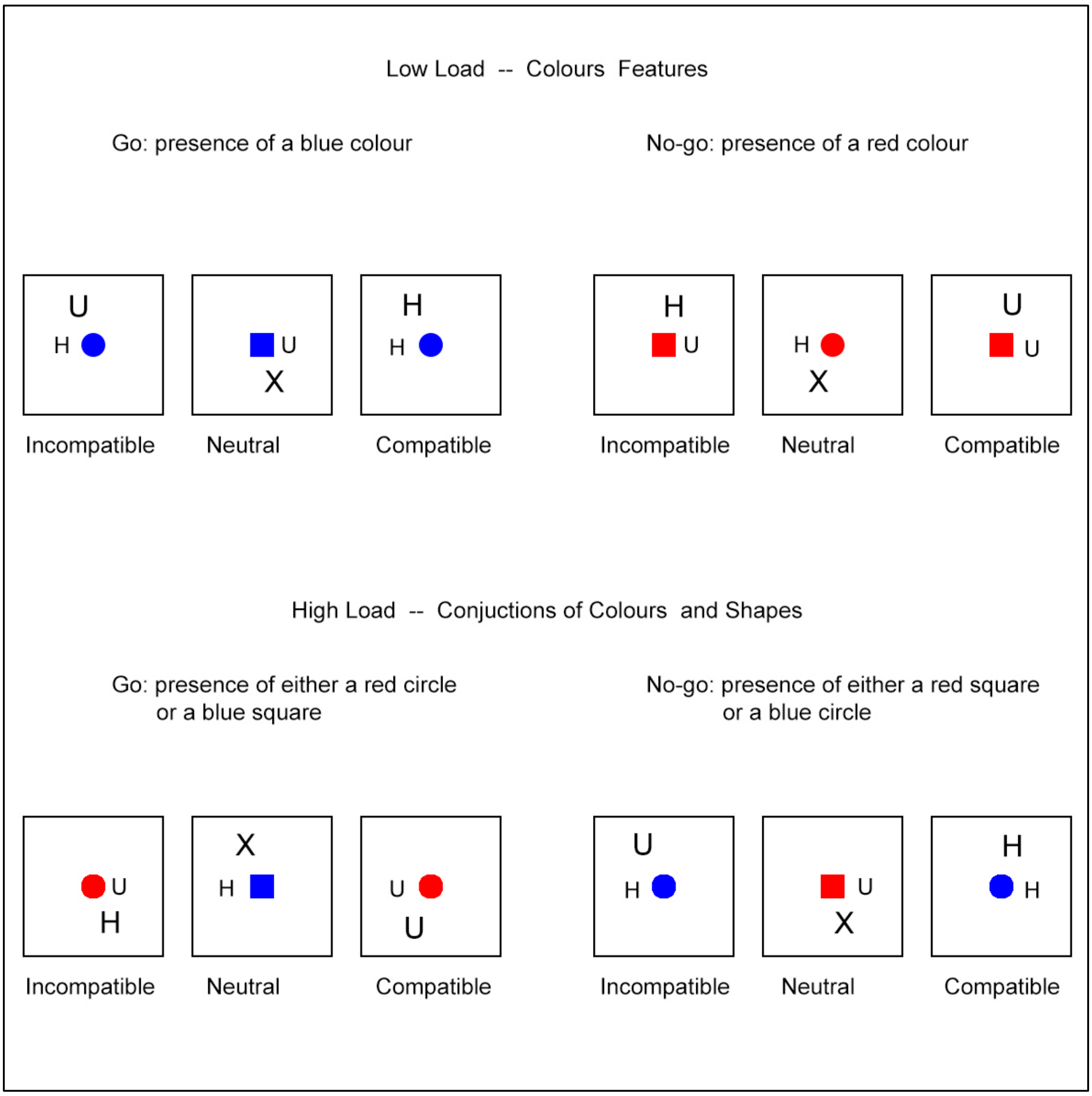

Figure 1.6. Examples of the manipulation of perceptual load by adaptation from feature integration theory. Adapted from Lavie (1995).

1.5.1.2 Effects of Working Memory Load on Irrelevant Distractor Perception

The studies of perceptual load lent support to the mechanism of early selective attention that prevents the processing of irrelevant distractors. This is the passive form of selective attention: distractor effects are increased in the situations of low perceptual load, compared to those in the situations of high perceptual load. Nevertheless, the active form of selective attention (the late-selection mechanism) plays the opposite role in preventing the processing of irrelevant distractors: distractors effects are increased under high working memory load and are decreased under low working memory load. The left example in Figure 1.7 illustrates the study conducted by Lavie, Hirst, de Fockert, and Viding (2004). Each trial began with a memory set in which working memory load was manipulated by the size of memory set. In the low working memory load, participants were asked to maintain only one digit, while in the high working memory load, they were presented with six digits from 1 to 9 in a random order and were asked to remember the six digits. Under low and high working memory conditions, a target letter (x or z) and a peripheral irrelevant item were shown to participants. The irrelevant item was compatible (an *X* when the target was x), incompatible (an *X* when the target was z), or neutral (the letter *N*) with the target letter. Lastly, a one-digit memory probe was presented on the screen. After indicating whether the target letter was x or z in the selective attention task, participants needed to indicate whether the memory probe was present or not in the memory set. The results supported the prediction that increasing working memory load induced greater distractor effects from task-irrelevant distractors, with a larger distractor effect of 193 ms under high working memory load than that of 140 ms under low working memory load. The findings were consistent with load theory which predicted when working memory is overloaded, the ability to inhibit the processing of task-irrelevant stimuli is weakened, and thus the distractor effects from task-irrelevant stimuli are increased. The results provided evidence to the cognitive mechanism hypothesis.

This cognitive mechanism hypothesis was also demonstrated in another study (de Fockert, Rees, Frith, & Lavie, 2001). The procedure was illustrated in the right example of Figure 1.7. A set of digits was shown to participants. Under low working memory load, the digits were 0, 1, 2, 3, and 4, which were always shown in the fixed ascending order, that is, 01234; under high working memory load, the same digits were presented, but shown in a different order on each trial (e.g., 02341 or 04213). Participants were asked to memorize not only the five digits but also the order of the five digits. Then, a face image, on which a name was superimposed, was presented to participants. The identities of distractor faces were either congruent with the written names, or incongruent with the written names, or anonymous. Finally, a memory probe was shown to participants. In the selective attention task, participants were instructed to categorise the written names as pop stars or politicians while ignoring irrelevant faces. In the memory task, they were asked to indicate the digit that followed the memory probe in the memory set (to press “4” in the example below). The results showed that reaction times were longer under high working memory load than under low working memory load. The activation of the fusiform gyrus and extrastriate visual cortex related to the presence of the task-irrelevant faces was greater under high memory working load than under low memory working load. Together, the results from behavioural and functional imaging studies demonstrated that the ability to ignore irrelevant faces was impaired when working memory load was increased, thus eliciting greater distractor effects from task-irrelevant faces.

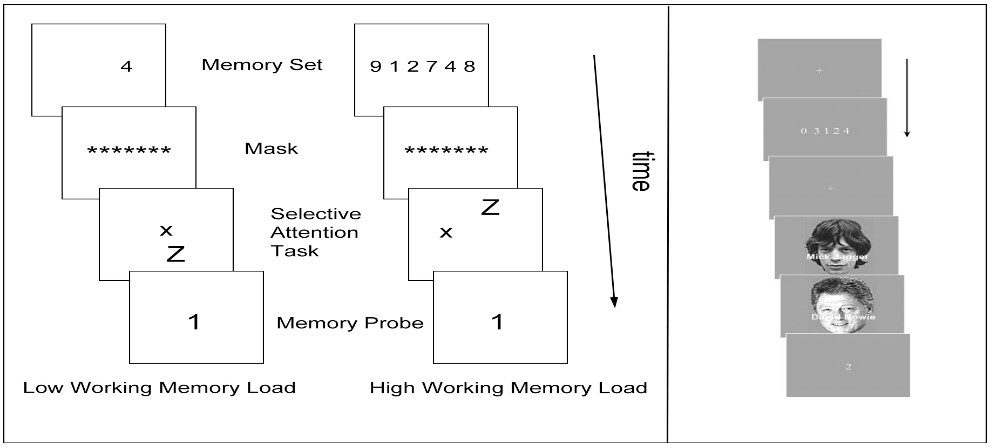

Figure 1.7. Examples of the manipulation of working memory load. The left example is the stimuli in the study of Lavie et al. (2004). The right example is the stimuli in the study of de Fockert et al., (2001). Adapted from Lavie et al. (2004) and de Focker et al. (2001).

However, recent studies (Kim, Kim, & Chun, 2005; Oh & Kim, 2004) have challenged this claim, suggesting that the impairment to selective attention caused by working memory load depends on the type of working memory task. Oh and Kim (2004) found that the spatial working memory task impaired the performance of visual search, while the non-spatial, colour working memory task did not interfere with the visual search processing. The implication of their study was that if working memory items and targets compete for the same processing resources, greater distractor effects from irrelevant distractors should occur under the conditions of high working memory load, whereas if the processing resources required for working memory items overlap with those for task-irrelevant distractors, no distractor effects from irrelevant distractors will be expected under the high working memory load conditions. Park, Kim, and Chun (2007) found that when maintaining faces in working memory, the distractor effects from irrelevant distractors (houses) on the face-matching task were increased. By contrast, when maintaining houses in working memory, the distractor effects from irrelevant distractors (houses) on the face-matching task were decreased, and the processing of the face-matching task was even found to be facilitated. The findings implied that the effects of working memory load on the selective attention task (impairment or benefit) rely on whether the type of working memory load overlaps with the processing resources required for the target or distractor processing.

In sum, the empirical data provided evidence to support load theory. Under high perceptual load, the increased demands of task-relevant stimuli could eliminate or reduce distractor effects from task-irrelevant distractors, which results in the early selection. In contrast, under low perceptual load, any remaining attentional resources not occupied by task-relevant stimuli will be involuntarily allocated to task-irrelevant stimuli to induce larger distractor effects, which results in the late selection.

1.5.2 A Summary of Early and Late Event-Related Potential Components Associated with Facial Expression Processing

It is commonly claimed that processing occurring at early stages is mandatory automatic in nature, which does not need attentional resources (Meeren et al., 2005; Palermo & Rhodes, 2007), whereas a processing happening at late stages requires attentional resources (Gu, Mai, & Luo, 2013). ERP techniques have been often used to investigate the time course of neuronal processing with a high temporal resolution at the millisecond level. The technique of ERPs measures electrical activity of the brain by using electroencephalography (EEG) scalp recordings after the onset of a stimulus. ERPs are voltage deflections that can be elicited by a variety of sensory, cognitive, or motor events (Sur & Sinha, 2009). The ERPs waveforms (Figure 1.8) are comprised of a series of positive-going (P) and negative-going (N) deflections that are termed "waves" or "components". The numbers following the "P" or "N" reflect the temporal sequence of their appearance. It is suggested that the latency of ERP components can be used to index the time course of cognitive processing (Luck & Hillyard, 2000). Some earlier (< 300 ms) and later (> 300ms) ERP components in regard to the processing of emotional stimuli will be introduced in the following section.

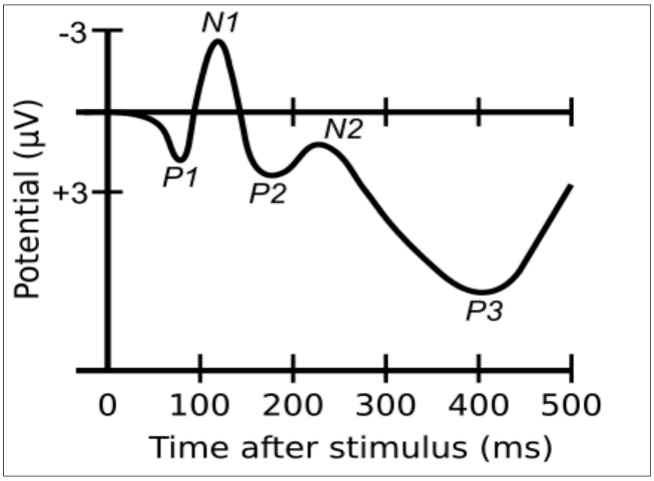

Figure 1.8. A waveform with some major ERP components.

1.5.2.1 Early Visual Components

The **P1** component, also called P100, is the first positive-going wave with a peak at around 100-130 ms (Luck, 2005). It is reported that the P1 component is sensitive to physical stimulus factors such as contrast (Olofsson, Nordin, Sequeira, & Polich, 2008) and is an index of early perceptual processing. It is found that the amplitude of P1 is more positive for faces than non-face objects (Itier & Taylor, 2004; Herrmann, Ehlis, Ellgring, & Fallgatter, 2005). Additionally, there is evidence that the P1 component is sensitive to affective picture stimuli, being larger for unpleasant than pleasant target stimuli (Carretie, Hinojosa, Martin-Loeches, Mercado, & Tapia, 2004a; Carretie, Mercado, Hinojosa, Martin-Loeches, Sotillo, 2004b).

N170 is a negative-going component, occurring between 120 and 220 ms and peaking around 170 ms after the onset of stimuli. There is a consensus that the N170 is linked with the structural encoding of faces, with larger amplitude and longer latency for faces upside-down than faces normally presented (Watanabe, Kakigi, & Puce, 2003; Stekelenburg, & de Gelder, 2004). Besides, it was observed that the N170 might be related to emotional information shown on the faces. Some studies have reported that the N170 was not modulated by facial expressions (Eimer & Holmes, 2002; Holmes, Vuilleumier, & Eimer, 2003), whereas other studies have found that the N170 amplitude was sensitive to emotional facial expressions (Batty & Taylor, 2003; Miyoshi, Katayama, Morotomi, 2004), at least a larger amplitude for fearful faces compared to neutral faces (Batty & Taylor, 2003). Recently, Thierry, Martin, Downing, and Pegna (2007) have argued against the consensus of the face-specificity of the N170, revealing that the N170 was modulated by low-level visual features (e.g., inter-stimulus perceptual variance), not by visual object categories (faces vs non-faces). Therefore, in order to figure out whether the N170 is sensitive to emotional facial expressions, further studies should be carried out with careful manipulation of low-level visual features of stimuli.

1.5.2.2 Early Auditory Components

N1 or N100 is a negative-going evoked potential, peaking between 80 and 120 ms after stimulus onset. The N1 is related to the perception of auditory stimuli. The amplitude of N1 is influenced by the rise time of stimulus onset (Spreng, 1980), as well as some other factors, such as inter-stimulus interval (ISI) (Davis, Mast, Yoshie, & Zerlin, 1966), stimulus intensity (Keidel & Spreng, 1965), arousal (Nash & Williams, 1982), and auditory selective attention (Hillyard, Hink, Schwent, & Picton, 1973). It has also been reported that the N1 went more negative for neutral vocalisations than for emotional ones (Liu, Pinheiro, Deng, Nestor, & McCarley, 2012).

P2 (also known as P200) wave is a positive-going component, following the N100 component. The P2 is responsive to the onset of auditory stimuli, peaking at around 150 to 250 ms. P2 is modulated by feature detection (Luck & Hillyard, 1994), selective attention (Hackley, Woldorff, & Hillyard, 1990), and other early sensory stages of item encoding (Dunn, Dunn, Languis, & Andrews, 1998). P2 is related to perceptual processing of emotional auditory stimuli, with one study indicating the P2 amplitude was greater for happy than for sad voice stimuli (Spreckelmeyer, Kutas, Urbach, Altenmuller, & Munte, 2006), and another study finding the enhanced P200 for neutral prosody as compared to for emotional prosody (Liu et al., 2012).

1.5.2.3 Late Visual Components

N2 (also called N200) is a negative-going wave concerning cognitive control processing. The N200 peaks around 200-350 ms after stimulus onset and is found over anterior scalp sites (Folstein & van Petten, 2008). It has been found that stimulus valence modulates the N2 wave, with pleasant stimuli eliciting the larger N2 compared with unpleasant stimuli (Carretie et al., 2004a; 2004b).

P3 (P300) is the most studied later ERP component. The P3 wave is a positive deflection with the latency between 250 and 500 ms (Polich, 2007). The P300 is suggested to be associated with stimulus evaluation (Willadsen-Jensen & Ito, 2006) and to be sensitive to response conflict (Bartholow et al., 2005; Folstein & Van Petten, 2008). Recently, a study (Wang et al., 2012) found that the P3 was enhanced under the condition of conflicting emotional information, such as when attended stimuli were fearful faces, while unattended stimuli were neutral faces. Another affective ERP study (Delplanque, Lavoie, Hot, Silvert, & Sequeira, 2004) has found that pleasant stimuli triggered the greater P3 amplitudes than unpleasant stimuli did.

The following subsections will review and discuss behavioural and ERP studies concerning to the issue whether attentional resources influence context effects on recognition of facial expressions.

1.5.4 Visual context effects and attentional resources

A limited number of studies explored whether the effects of affective visual contextual information (such as bodily expression, emotional scene) on recognition of facial expressions are independent of attentional resources.

Righart and de Gelder (2008b) investigated whether the recognition of facial expressions was influenced by contextual stimuli when the task load was increased using a letter-recall task. They placed an “x” or “o” on the centre of each face that was superimposed on scene images. Participants were asked to categorise the emotion of the faces (fear, disgust, and happiness), then they needed to indicate the letter on the face. Righart and de Gelder stated that according to load theory, if the effects of emotional scenes on the recognition of facial expressions need attentional resources, context effects will diminish under high task loads. Therefore, it was concluded that context effects are free of attentional resources, which was reflected by the results that context effects on facial expression recognition were exist when performing the letter-recall task. However, the type of task loads might be confused, as perceptual load and memory load exert different influences on task-irrelevant distractors (in this study, irrelevant distractors were the emotional scenes). This experiment used the letter-recall task that was a memory task to consume attentional resources. Load theory proposed that high perceptual load leads to the decrease of distractor effects from task-irrelevant distractors, whereas high memory load induces the increase of distractor effects from task-irrelevant distractors. So, context effects (that is, distractor effects) caused by emotional scenes (irrelevant distractors in this experiment) on recognition of emotional faces will not be decreased under memory task load, but will be potentially increased. Unfortunately, it remains unknown whether context effects were increased, as this study did not compare context effects between the conditions of low and high memory task load. Following the logic of load theory, the results of this study could not be interpreted as that context effects do not depend on the attentional resources.

Similar evidence comes from the study of Aviezer et al., (2011) in which they used a memory task to manipulate the task (working memory) load to test whether the context effects are free of attention or not. They presented a letter-number string including five numbers and one letter (e.g., 274K81) to participants. For the high cognitive load group, participants were asked to memorize the whole string in the correct order, while for the low cognitive load group, participants were asked to memorize the only letter in the string. After the presentation of the string, participants were shown a contextualized facial expression (a face-body stimulus) and were asked to categorise the facial expression. The last thing that they were required to do was to indicate which string was seen previously on the screen. They got feedback after each trial for the memory task. The results were that the patterns of context effects on recognition of facial expressions were similar under low and high cognitive loads, suggesting that context effects were attention-free. As predicted by load theory, high working memory load should result in context effects from irrelevant bodily expressions on recognition of facial expressions to be increased rather than to be decreased (Lavie et al., 2004; de Fockert et al., 2001). Some empirical studies (Kim et al., 2005; Oh & Kim, 2004; Burnham, Sabia, Langan, 2013) indicated that it is the type of working memory load that can determine whether distractor effects from irrelevant distractors on primary tasks would be increased or decreased. When processing resources for working memory load neither overlap with those for a primary target task nor with those for an irrelevant distractor, performance of the primary task might be unaffected by the irrelevant distractor. Due to uncertainty of the impact of working memory load on irrelevant distractors, it might be inappropriate to conclude that context effects on recognition of emotional facial expressions are free of attentional resources.

Additionally, ERP was used to investigate whether context effects on the recognition of emotional facial expressions occurs at the early stage of processing or at the late stage of processing. Meeren et al. (2005) found that the P1 amplitudes were influenced by the factor of congruency, with the larger P1 amplitudes for incongruent face-body stimuli than for congruent face-body stimuli, which indicated that context effects on recognition of emotional facial expressions took place at the early stage. Therefore, Meeren et al. (2005) suggested that context effects on recognition of emotional faces do not depend on attentional resources. Similar results were found in another study (Righart & de Gelder, 2008a) where the N170 amplitudes went more negative when faces were in fearful scenes than in happy and neutral scenes, implying that the concurrently presented irrelevant scenes influence recognition of emotional facial expressions during the early stage of processing, without the requirement of attentional resources.

Gu et al. (2013) conducted a study to further investigate at which stage context effects from irrelevant bodily expressions on recognition of facial expressions. The procedure for this study was similar to the two studies above (Meeren et al., 2005; Righart & de Gelder, 2008b), in which the face-body stimuli, either congruent (happy face-happy body and fearful face-fearful body) or incongruent (happy face-fearful body, and fearful face-happy body) were presented to participants one at a time. The task required participants to press F or J on the keyboard to answer which facial expression it was, ignoring the accompanied bodily expression. The ERP data showed that the P1 waves were larger when faces were accompanied by fearful bodies than when the same faces were accompanied by happy bodies, regardless of the emotion expressed by the faces. The N2 components and the P3 components were both sensitive to congruency effects. The difference between these two components was that the larger N2 amplitudes were elicited by incongruent face-body stimuli than by congruent face-body stimuli, while the P3 waves went more positively for faces with congruent bodily expressions than for faces with incongruent bodily expressions. The findings suggested that context effects on recognition of emotional facial expressions begin at the early processing stage, which is consistent with the results in the previous studies (Meeren et al., 2005; Righart & de Gelder, 2008b). What was new in this study was that context effects extended to the later evaluative processing stage, reflected by the larger P3 amplitudes caused by congruent compound stimuli. The results suggested that attentional resources might be required for context effects on the recognition of facial expressions.

Researchers (Meeren et al., 2005; Righart & de Gelder, 2008b) reported that context effects proceeded without the need of attentional resources, as it happens at the early stage of processing. However, attention can operate either at the early stage under some conditions (e.g., high perceptual load) or at the late stage (Luck, Woodman, & Vogel, 2000). As noted earlier, it has been proposed that under the condition of high perceptual load, attentional resources take effect at the early stage, which means the processing of irrelevant contextual emotional information is suppressed, and then will not induce any effect on the recognition of emotional facial expressions. In situations of low perceptual load, attentional resources are not fully occupied by the relevant task, and the remaining capacity will be allocated to task-irrelevant contextual emotional information, resulting in interference effects on the relevant facial expression recognition task (Lavie, 1995, 2005). Thus, the evidence from the two ERP studies (Meeren et al., 2005; Righart & de Gelder, 2008b) that context effects are free of attention is not decisive, unless context effects happened at the early stage under high perceptual load. Moreover, according to the results in Gu et al.'s study (2013), the conflicting signals from emotional faces and bodies were detected early in temporal course, and then were resolved late in time course, implicating attentional resources might modulate context effects on recognition of facial expressions.

1.5.5 Auditory Context Effects and Attentional Resources

The terms multisensory integration and cross modal interaction are often used interchangeably. Nonetheless, there is difference between these two terms. Multisensory integration is used to describe a process in which information from different sensory modalities is combined together to yield a coherent percept. (Spence, Senkowski, & Röder, 2009). Cross modal interaction refers to the influence of a stimulus in one modality on perceiving a stimulus in another modality (Spence et al., 2009).

One of the most commonly mentioned examples of human multisensory integration is speech perception: simultaneously hearing a word and seeing lip movement. Matched visual and auditory information can improve and enhance the perception of speech, with higher accuracy and faster reaction times (Mozolic, Hugenschmidt, Peiffer, & Laurienti, 2008). Mismatched information can influence the speech process, which may result in the McGurk effect (McGurk and MacDonald, 1976). They found that hearing an auditory syllable (i.e., /ba/) and seeing a talker articulating another syllable (i.e., /ga/), led to the syllable being perceived as “tha” or “da”, which was distinct from either the auditory or the visual signal.

For the last 30 years or so, it has been widely investigated whether the multisensory integration depends on attentional resources. Recent findings, challenging the prevalent view that multisensory integration is an automatic process and is not affected by attention (Bertelson, Vroomen, De Gelder, & Driver, 2000; Driver, 1996; Foxe et al., 2000; Vroomen, Bertelson, de Gelder, 2001a, 2001b), revealed that attentional resources are required in the process of multisensory integration in behavioral and neuroscientific studies (Busse, Roberts, Crist, Weissman, & Woldorff, 2005; Fairhall & Macaluso, 2009; Talsma & Woldorff, 2005; Talsma ,Doty & Woldorff, 2007).

The influences from auditory emotional stimuli on recognition of emotional facial expressions are an example of cross-modal interaction. A few studies attempted to discover whether attention plays a role in the auditory context effects on the recognition of emotional facial expressions. So far, there is no consensus, which is similar to the situation for multisensory integration. The followings detail the studies pertaining to the role of attention in cross-modal interaction of emotional information from visual and auditory modalities.

Vroomen, Driver, and de Gelder (2001c) examined this topic and suggested that the influence of facial expressions on emotional voice was independent of attentional resources. They presented an emotional facial expression (happy or fearful), simultaneously with an emotional vocal stimulus to participants in their three experiments. The first two experiments adopted additional visual demanding tasks, an add task in Experiment 1 (a task to report the sum of the two visual digits between 1 and 9 shown) and a count task in Experiment 2 (a task to detect and count the number of zeros in a rapid serial visual presentation), while the third experiment used an demanding auditory task (judging the pitch of a tone was high (540-Hz) or low (500-Hz) in Experiment 3). Participants were asked to indicate whether the emotion from auditory modality was happy or fearful, and to ignore the accompanied facial expression. At last, they had to complete the additional task. The results of the three experiments showed that the impact of emotional faces on perception of emotional voices was not affected when completing the additional demanding task. Accordingly, it was inferred that the attentional resources are not recruited in the cross-modal interaction of visual and auditory emotional information.

Further evidence comes from an ERP study (Pourtois, de Gelder, Vroomen, Rossion, & Crommelinck, 2000). Participants were shown face-voice stimuli one at a time, either a congruent pair (an angry face/an angry voice or a sad face/a sad voice) or an incongruent pair (an angry face/a sad voice or a sad face/an angry voice). This experiment was a passive-viewing oddball task. Participants were only instructed to pay attention to emotional facial expressions and to ignore auditory stimuli. Electrophysiological data were recorded and the waveforms were analysed by comparing the congruent and incongruent trials. The larger N1 amplitudes were found in the congruent trials, and the smaller P2 amplitudes were observed in the congruent pairs. It was suggested that the influences of emotional voices on recognition of facial expressions are a mandatory, automatic process, which is considered to be a process without the need of attentional resources. Thus, the results indirectly showed that the crossmodal interaction of emotional perception do not depend on attentional resources.

However, another study (de Jong, Hodiamont, & de Gelder, 2010) challenged the idea that the cross-modal interaction of emotional faces and voices is free of attention. de Jong et al. (2010) demonstrated that modality-specific attention affected cross-modal interaction of visual and auditory affective information. In the first task, the target emotion was happy or sad; in the second task, the target emotion was happy or fearful. In both tasks, participants were presented with emotional auditory stimuli, simultaneously with congruent or incongruent emotional facial expressions. Additional visual and auditory tasks were used in each of the two tasks. The additional visual task (a digit of "6" or "8") was placed between the eyebrows; the additional auditory task was comprised of two pairs of tones. For one pair, each tone had the same frequency of 500 Hz; for the other pair, one tone had a frequency of 500 Hz; the other tone had a frequency of 540 Hz. Participants were instructed to categorise the emotion of the sound, while ignoring the faces. And then they were asked to respond to the question "was there an 8 in the face?" when facing the visual demanding task, or to respond to the question "did you hear a higher tone?" when for the auditory demanding task. In both tasks, visual and auditory demanding tasks attenuated the cross-modal interaction; that is, the influence of emotional faces on the categorisation of emotional voices was decreased. Therefore, it was suggested that visual influences of facial expressions on perception of auditory emotional voices require attentional resources.

Two behavioural studies described in this section were to explore the role of attention in the impact of visual emotional faces on recognition of auditory emotional stimuli, rather than whether the influence of auditory stimuli on recognition of facial expressions relies on attentional resources or not. It has been reported that the visual effects on processing of auditory stimuli might not be processed in the same way as the auditory effects on processing of visual stimuli (Talsma et al., 2007). Collignon et al. (2008)

found that irrelevant facial expressions exerted more power on the recognition of emotional voices than irrelevant emotional voices on perception of emotional faces. Hence, the evidence on whether attention is required in context effects of facial expressions on recognition of auditory emotional stimuli cannot be used to make a conclusion on whether context effects of emotional voices on recognition of facial expressions is independent of attentional resources or not. In sum, current evidence is mixed with regard to the question whether the auditory contextual effects on recognition of facial expressions requires attentional resources.

1.5.6 Summary

Regardless of visual context effects or auditory context effects on perception of facial expressions, there is no conclusive study on the topic of the relationship between attention and context effect. In the light of the mixed situation, more research should be carried out in the future.

1.6 Overview of the Current Experiments

Before testing the main hypothesis, the material used as context stimuli were evaluated in chapter 3, i.e. Bimodal Face and Body Gesture Database (Gunes & Piccardi, 2006) (Experiment A) and Montreal Affective Voice (Belin, Fillion-Bilodeau, & Gosselin, 2008) (Experiment B). The aims of Chapter 3 were whether bodily expressions and vocal expressions are perceived as intended target emotions so that they could be used as appropriate contextual stimuli in the following main test.

As reviewed above, a number of studies have found that the recognition of facial expressions is influenced by emotional contextual information. Inconsistent patterns of context effects were observed in previous studies. For example, the recognition of happy faces benefits from congruent contextual information, compared to incongruent one (de Gelder & Vroomen, 2000; van de Riet & de Gelder, 2008), whereas other studies did not find significant differences between happy faces paired with happy context and happy faces paired with incongruent context. The situation for fearful faces is similar to that for happy faces. Emotion seeds view (Aviezer et al., 2008a, 2008b) seems to explain different patterns of context effects observed in previous studies. Emotion seeds view hypothesizes that context effects are greater when faces are paired with 'high similarity context' than when faces are paired with 'low similarity context' and participants are more likely to mis-categorise emotional faces as the emotion conveyed by high similarity contexts than to mis-categorise emotional faces as the emotion conveyed by low similarity contexts. Chapter 4 examined whether the pattern of context effects proposed by emotion seeds view could apply to all six basic facial expressions.

It has been found that the magnitude of context effects is larger for a type A ambiguous facial expression than for an unambiguous facial expression and is larger for a more ambiguous mixed facial expression (e.g., 50%happiness - 50%sadness) than for a less ambiguous one (e.g., 90%happiness - 10%sadness) (de Gelder & Vroomen, 2000; van den Stock et al., 2007). Chapter 5 explored the relationship between contextual information (bodily expressions and vocal expressions) and the recognition of different levels of intensity of facial expression. The aim of these two studies was to ascertain whether the magnitude of context effects for different levels of intensity of facial expression is similar to that for Type A ambiguous facial expressions. That is, whether the magnitude of context effects is smaller for unambiguous emotional faces than for lower levels of intensity.

As previously mentioned, there is no conclusion on the issue of whether context effects are influenced by attentional resources or not. Chapter 6 was designed to investigate this issue under the framework of perceptual load. The aim of the last three experiments in chapter 6 was to find out whether the magnitude of context effects will be reduced under high perceptual load. If so, as proposed by load theory, context effects will be influenced by attentional resources. If not, context effect will not be influenced by attentional resources.

Chapter 2 General Methods

This chapter describes research methods that were common to all the experiments. Specific details differing between experiments will be provided in relevant chapters.

2.1 Participants

All participants recruited volunteered to take part in the experiments. They were paid appropriately for their time. All had given informed consent. All reported normal or corrected to normal vision, and participants who took part in the experiment with auditory stimuli reported normal hearing. The study was performed in accordance with the ethical standards of the Massey University Human Ethnics guidelines and was deemed to be low risk.

2.2 Apparatus and Stimuli

The experiments were programmed in E-prime 2.0 and run on a computer running Windows XP. All visual stimuli were presented on a 24-inch View Sonic monitor with a resolution of 1920 x 1080 pixels and with a 60 Hz refresh rate. All of the experiments were conducted in an illuminated room. And the auditory stimuli were presented via SENIC IS-R1 headphone at self-adjusted comfortable level.

The images of facial expressions were taken from Karolinska Directed Emotional Faces (KDEF) set (Lundqvist, Flykt, & Öhman, 1998). The images of bodily expressions came from Bimodal Face and Body Gesture Database (FABO) (Gunes & Piccardi, 2006). The auditory stimuli were chosen from Montreal Affective Voices (MAV) set (Belin et al., 2008). The selection and manipulation of research material are described in the relevant chapter.

2.3 Procedure

A typical trial is illustrated in Figure 2.1. Each trial consisted of the following sequence. A fixation cross was displayed for 500 ms, followed by a stimulus (including a face-context stimulus and an isolated face) that was presented for 1830 ms. To keep the consistency between studies of visual stimuli and auditory stimuli, the choice of the duration (1830 ms) came from the longest auditory stimuli (e.g., laughter). Then a response screen with six options (happiness, surprise, fear, disgust, sadness, anger) was shown on the screen until participants used mouse to click on one of the choices or it disappeared after 8 s without any response, in chapter 3, chapter 4. Finally, a blank screen was shown for 1 s. Feedback was provided to participants in the last three experiments, whereas no feedback in the other experiments. Each stimulus was presented 10 times in a random order. The response in chapter 5 and 6 was different from that in chapter 3, and 4 and will be introduced in chapter 5 and 6.

Participants completed practice trials before the main test. After each task was completed participants were offered debriefing and were thanked for their participation.

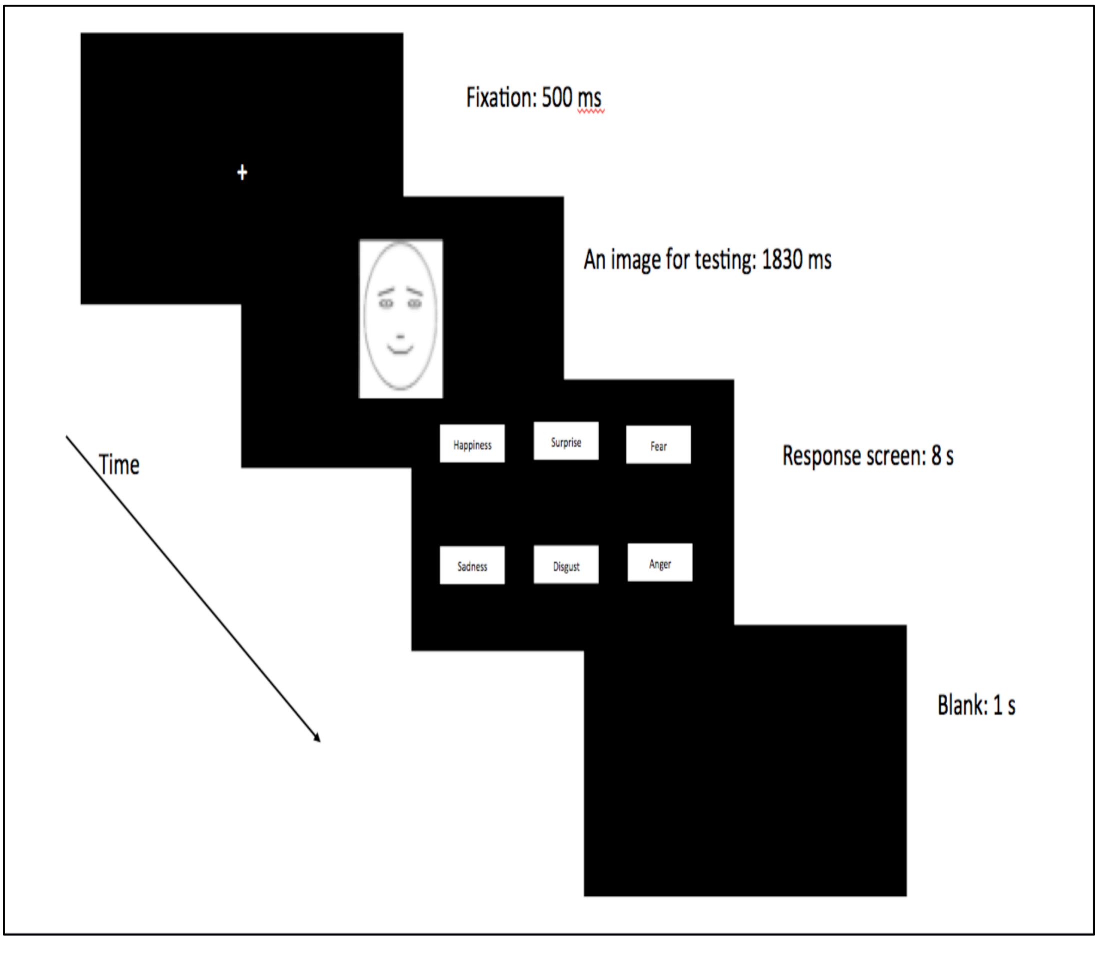

Figure 2.1. An illustration of a typical trial. A 500 ms fixation cross is followed by an image that was presented for 1830 ms. The response screen with six options of happiness, surprise, disgust, sadness, fear, and anger was presented for 8 s. At last, a blank screen was shown for 1 s.

Chapter 3 Evaluation of Research Material as Context Stimuli

Before conducting the following main studies, context stimuli were carefully chosen and evaluated. This chapter explored whether the research stimuli chosen for context stimuli are perceived as intended emotions.

3.1 Evaluation of Bodily Expressions (Experiment A)

A number of bodily expression databases provide convenience to studies on bodily expressions. As mentioned above, emotional bodily expressions may provide reliable emotional cues. Compared to facial expressions, bodily expressions have their own advantage for perceiving emotions as they can be recognised far away from the observers (de Gelder, 2009).

Studies of facial expressions generally agree on the fact that happy faces are recognised faster and more accurately than other emotional faces (Feyereisen et al., 1986; Harwood, Hall, & Shinkfield, 1999; Kirita & Endo, 1995; Kirouac & Doré, 1983). However, the results for recognition of happy bodies were mixed, where one study found happy bodies were the most difficult to be recognised (de Gelder & van den Stock, 2011), while other studies did not (Atkinson et al., 2004; Schindler et al., 2008; Thoma, bauser, & Suchan, 2013). Regardless of the mixed results, these mentioned studies have found that emotional bodily expressions were well recognised, reflected by high percentages of accuracy of recognition of the bodily expressions (de Gelder & van den Stock, 2011; Thoma et al., 2013). These mentioned studies focused on the recognition of whole body expressions. It was reported that upper-body expressions contain a rich source of expressive information (e.g., hands with clenched fists representing anger) and are sufficient for people to recognise emotion (Glowinski et al., 2008; Roether, Omlor, & Giese, 2010; Wallbott, 1998). For the current study, visual contextual stimuli used in the following main experiments are upper-body expressions which were selected from FABO set. However, previous studies adopting FABO set were concentrated on the aspect of emotional feature extraction (Chen, Tian, Liu, & Metaxas, 2011, 2013; Shan, gong, McOwan, & Peter, 2007) rather than on the validation of this database. The purpose of this experiment was to explore whether the chosen images of bodily expressions are perceived as intended emotions, thus they could be used as reliable cues for conveying emotion.

3.1.1 Method

3.1.1.1Participants

Fifteen participants, ranging from 19 to 25 years, took part in the experiment. Eight participants were female and seven were male. Eleven participants were right-handed and four were left-handed. All participants had given informed consent and were paid NZD 2 for participation.

3.1.1.2 Apparatus and Stimuli

The apparatus for this experiment is described in Chapter 2.

The images of bodily expressions come from Bimodal Face and Body Gesture Database (FABO) (Gunes & Piccardi, 2006). Gunes and Piccardi (2006) used two cameras, one for only recording video of faces and the other one for recording video of whole upper-body movement. The frames in each video was segmented into different stages, starting from neutral stage, then onset stage, to apex stage, following by offset stage, finally back to neutral stage. Gunes and Piccardi (2008) pointed out that apex frames of upper-body movement were at the maximum of its spatial position and found that they can provide effective cues for people to discriminate emotions. The results are consistent with other studies that suggested that emotions are delivered mostly through facial expressions and upper-body motion (Ballihi, Lablack, Amor, Bilasco, Daoudi, 2015; Glowinski et al., 2008), and argued that it is sufficient for us to perceive emotions from upper-body motion (Volkova, Mohler, Dodds, Tesch, & Bülthoff, 2014). So, for the present study, bodily expressions were chosen from apex frames of upper-body video as visual contextual images.

After completing recording of the database, Gunes and Piccardi (2006) have six participants, who did not take part in recording, label each video in the database. The bodily expression images of two females and one male were chosen based on the original evaluation results, with four participants out of 6 labeling a video as intended emotion (e.g., four participants out of six labeled a happy bodily expression as happy). And these three actors posed all six basic expressions. The images of one female were selected for practice trials, and the remaining one female and one male were selected for testing in the main trials. The colour body images were converted into grey-scale pictures. The faces in body images were blurred (as shown in Figure 3.1), using the function of Gaussian Blur with radius of 15 pixels in Adobe Photoshop CS5 (trial version). The framed upper-body images were measured horizontally 30 cm and vertically 19 cm.

3.1.1.3 Design and Procedure

A one-way (emotion: anger, disgust, fear, happiness, sadness, or surprise) within-subjects design was used to evaluate the chosen bodily expressions, that is, whether they could be perceived as intended emotions to become reliable contextual stimuli.

The procedure for this experiment was the same as that described in Chapter 2. In this experiment, participants were asked to indicate the emotion conveyed by images of bodily expressions. They used the mouse to click on one of the six options provided to make a judgment. The position of the six options was balanced across participants.

Each stimulus was presented 10 times in a random order. In total, 120 trials were tested in this experiment. Before doing the test, they were asked to complete 12 practice trials. The images used in practice trials did not appear in the test.

Figure 3.1. Examples of Bodily Expression Images Tested in the experiment. The left one represents surprised bodily expression; the right one represents angry bodily expression.

3.1.2 Results

Table 3.1 shows the behavioural performance on recognition of the bodily expressions. The percentage of accuracy for judging bodily expressions was, on average, 80.67% and the accuracy rates from highest to lowest were Happiness (97.00%), Anger (96.00%), Surprise (86.00%), Sadness (70.33%), Disgust (68.67%), and Fear (66.00%) (Chance performance = 16.7%). The data were analysed using a one-way (emotion: anger, disgust, fear, happiness, sadness, or surprise) repeated ANOVA, revealing a main effect of emotion for recognising bodily expressions, F (2.49, 34.88) = 7.701, p = .001, η^2_p = .355. p value was corrected by Greenhouse-Geisser Epsilon when applicable. Fisher's least significant difference[2] (LSD) post-hoc analyses also confirmed significant differences between 6 emotional bodily expressions. Happy and angry bodily expressions were better recognised, without any significant difference between them, p = .531. The accuracy of happy bodily expressions was significantly higher than fearful facial expressions (p < .001), sad ones (p = .001), disgusted ones (p = .006), and surprised one (p = .034). Angry bodily expressions were more accurately recognised than disgusted ones (p = .011), fearful ones (p < .001), sad ones (p = .001) and were recognised marginally significantly different from surprised ones (p = .056). The recognition of surprised bodily expressions was recognised better than that of fearful ones, p = .009, and was not significant different from that of disgusted and sad ones, both ps > .05. Disgusted, fearful, and sad bodily expressions were recognised the worst, and were without significant differences between each pairs of the three emotions, all ps > .10.

---

[2]Fisher's least significant difference (LSD): It has been reported that Bonferroni adjustments are too conservative and increase type II errors (Armstrong, 2014; Perneger, 1998; Rothman, 1999; Saville, 1990; Sedgwick, 2014). So, post-hoc LSD analysis was adopted for all the experiments in this thesis.

Table.3.1. Mean Percentage of Accurate Recognition of Bodily Expressions

Bodily expressions	Accuracy (%)
Happiness	97.00 (0.82)
Anger	96.00 (1.48)
Surprise	86.00 (4.71)
Sadness	70.33 (5.86)
Disgust	68.67 (8.69)
Fear	66.00 (5.05)

- SEM in parentheses

3.1.3 Discussion

The main goal of Experiment A was to assess whether the bodily expressions chosen were perceived as target emotions. As anticipated, the images of bodily expressions were all recognised as intended emotions though some differences between the categories of emotions were significant. The results that the best accuracy rates for happy and angry bodily expressions are in parallel with the advantage for happy and angry facial expressions, showing higher accuracy and faster reaction times for detecting and perceiving happy and angry faces (Ekman & Friesen, 1971; Hasen & Hasen, 1988; Hess et al., 1997; Horstmann & Bauland, 2006). This experiment confirmed that upper-body expressions can clearly convey emotional information to people, which is consistent with previous studies suggesting bodily expressions are reliable emotional cues (de Gelder, 2009; Kleinsmith & Bianchi-Berthouze, 2013; Wallbott, 1998).

3.2 Evaluation of Vocal Expressions (Experiment B)

Another important source for emotion recognition comes from auditory modality, such as speech prosody, and nonverbal affect bursts. Speech prosody can convey affective information either from the tone of speech, or from the semantic information in speech (Belin, et al., 2008; Scherer, Ladd, & Silverman, 1984). Like speech prosody, nonverbal affect bursts, which are defined as "short, emotional non-speech expressions, comprising both clear non-speech sound (e.g., laughter) and interjections with a phonemic structure (e.g., 'Wow!'), but excluding 'verbal' interjections that can occur as a different part of speech (like 'Haven!' 'No!' etc.)" (Schröder, 2003, p. 103), could convey reliable emotional signals (Schröder, 2003). Affect bursts, with the advantage of no semantic information, minimized the interactions between affective processing and linguistic processing.

It has been found that some emotions are conveyed more naturally by affect bursts and are better recognised from affect bursts, whereas others are perceived more easily from speech prosody (Banse & Scherer, 1996; Scherer, 1981). For example, the accuracy for disgust expressed by speech was low at 15% (Banse & Scherer, 1996). In contrast, the accuracy for disgust expressed by affect bursts was high at 93% (Scherer, 2003). Conversely, anger was better recognised through speech than through affect bursts (Scherer, 2003).

However, compared to affective studies on speech prosody, a limited number of studies were carried out on affect bursts due to the lack of appropriate material. The database of Montreal Affective Voices (MAV) was set up for this reason and was chosen to be research material in this experiment. The purpose of this experiment was to explore whether the chosen nonverbal vocal expressions are perceived as intended emotions, thus they can be used as reliable emotional cues.

3.2.1 Method

3.2.1.1 Participants

Fourteen new participants, ranging from 18 to 30 years, took part in the experiment. Eight participants were female and six were male. All of the participants were right-handed. They all had given informed consent and were paid NZD 2 for their participation. All had normal or corrected-to-normal vision and normal hearing ability.

3.2.1.2 Apparatus and Stimuli

This experiment was run on the same computer as that described in chapter 2. The auditory stimuli were presented via SENIC IS-R1 headphone at self-adjusted comfortable level.

The auditory stimuli were chosen from MAV set (Belin et al., 2008). MAV set consists of 90 short, nonverbal affect bursts, delivering emotions of anger, disgust, fear, pain, sadness, surprise, happiness, and sensual pleasure (plus a neutral expression). The vocalisations were recorded in a soundproof room. The actors were instructed to pronounce French vowel *ah* (/a/, similar to the English *a* in *apple*) in different emotional tones.

Nonverbal affect bursts of 2 females and 1male were selected. As in experiment A, the auditory stimuli of 1 female were used in practice trials, and the remaining material was used in main trials.

3.2.1.3 Design and Procedure

A one-way (emotion: anger, disgust, fear, happiness, sadness, or surprise) within-subjects design was used to evaluate the chosen vocal expressions, that is, whether they can be perceived as intended emotions to become reliable auditory contextual stimuli.

The procedure for this experiment was the same as that described in Chapter 2, except that the images of bodily expressions were replaced by nonverbal affect bursts. In this experiment, participants were asked to indicate the emotion conveyed by vocal expressions. They used a mouse to click on one of the six options to make a judgment. The position of the six options was balanced across participants.

Each stimulus was presented 10 times in a random order. In total, 120 trials were tested in this experiment. Before doing the test, they were asked to complete 12 practice trials. The auditory material used in practice trials was not appeared in the test.

3.2.2 Results

Table 3.2 shows the behavioural data on recognition of the vocal expressions. The accuracy rate for judging vocal expressions was, on average, 88.27% and the accuracy rates from highest to lowest were Disgust (98.93%), Happiness (97.50%), Sadness (96.43%), Anger (85.36%), Surprise (82.14%), and Fear (69.29%) (Chance performance = 16.7%). A one-way repeated ANOVA revealed a main effect of emotion for recognising vocal expressions, F (2.34, 30.39) = 7.224, p = .002, η^2_p = .357. p value was corrected by Greenhouse-Geisser Epsilon when applicable. Pairwise comparisons

(Fisher's LSD) also confirmed significant differences of vocal expressions of six emotions. Happiness, disgust, and sadness were better recognised than fear and surprise, all *ps* $< .05$, except that the accuracy of sadness was marginally different from that of surprise, $p = .058$. There was no significant difference between each pair of the three vocal expressions (happiness, disgust, and sadness), all *ps* $> .10$. The accuracy of angry vocal expressions was marginally less than that of disgusted vocal expressions, $p = .083$, but was not significantly different from that of other vocal expressions, all *ps* > 0.10. Fearful and surprised vocal expressions were recognised the worst, with marginally difference between each other, $p = .054$.

Table 3.2. Mean Percentage of Accurate Recognition of Vocal Expressions

Vocal Expressions	Accuracy (%)
Disgust	98.93 (0.77)
Happiness	97.50 (1.46)
Sadness	96.43 (2.48)
Anger	85.36 (7.18)
Surprise	82.14 (6.46)
Fear	69.29 (4.09)

- SEM in parentheses

3.2.3 Discussion

The main goal of experiment B was to assess whether the vocal expressions chosen were perceived as target emotions. The average recognition accuracy for the six vocal expressions was 88.27%, which is similar to the results (81% for 10 emotion categories) observed in Schröder's study (2003). As expected, the nonverbal affect bursts selected were all perceived as intended emotions though some differences between the categories of emotions were significant.

The worst accuracy rate for fearful vocal expressions is in line with the lowest recognition accuracy observed for fearful bodily expressions. This result again demonstrated the fact that fear is the most difficult emotion for people to recognise (Biehl et al., 1997; Russell, 1994; Suzuki, Hoshino, Shigemasu, & Kawamura, 2006). The high accuracy for recognition of vocal expressions confirmed that nonverbal affect bursts are highly effective and reliable cues for emotion recognition (Russell, Bachorowski, & Fernández-Dols, 2003).

Chapter 4 Basic Facial Expressions in Contexts

In daily life, objects always appear with a background or other objects. There is a general consensus that when objects are presented in an appropriate scene, such as a plate in a kitchen, objects are detected and recognised easily and quickly, whereas objects presented in an inappropriate scene, such as a violin in a kitchen, recognition of objects becomes difficult and slow (Bar, 2004; Oliva & Torralba, 2007). In other words, a scene that is physically or semantically consistent with objects facilitates perception and recognition, at least in contrast with a scene that is inconsistent with objects.

Perception of facial expressions, like objects in context, has found to be influenced by concurrently presented contextual information. Most studies made the comparison between emotionally congruent face-context pairs and emotionally incongruent face-context pairs and drew a conclusion that context effects do occur when recognising emotions from facial expressions. Moreover, emotion seeds views (Aviezer et al., 2008a, 2008b) proposed that recognition of facial expressions is systematically influenced by emotional context that concurrently appears with emotional faces. It was suggested that the greatest magnitude of context effects occurs when a facial expression is paired with a highly perceptual similar context, the moderate magnitude of context effects occurs when a facial expression is paired with a less perceptually similar context, and the lowest magnitude of context effects occurs when a facial expression is paired with a least perceptually similar context. Emotion seeds views were demonstrated in the lab of Aviezer and coworkers (2008b), whereas other researchers (Boucher, 2014; Mondloch, 2012; Mondloch et al., 2013b) did not find the same pattern of context effects as predicted by emotion seeds view. Boucher (2014) attempted to test whether emotion seeds views are applicable to other basic emotions of facial expression. He used six basic facial expressions plus neutral faces and re-evaluated perceptual similarity between each pair of seven types of faces. The pattern of context effects observed in this study was not consistent with the hypothesis predicted by emotion seeds view. For example, it was supposed that fearful facial expressions were categorised as surprised in surprised context (high similarity) more frequently than fearful facial expressions were categorised as sad in sad context (medium similarity). On the contrary, sad context resulted in larger influence on perception of fearful facial expressions than surprised context did. Boucher (2014) suggested that emotion seeds view was not adequate to explain how contextual information affects recognition of emotional facial expressions. He argued that context effects observed in previous studies (Aviezer et al., 2008b; Aviezer et al., 2011) which demonstrated the hypothesis of emotion seeds view were elicited by combination of bodily expressions and paraphernalia (e.g. a dirty diaper or a tombstone), rather than by pure bodily expressions. It is noted that neutral faces were included to examine the hypothesis of emotion seeds view. Nonetheless, the hypothesis was originally proposed for emotional facial expressions, not including neutral faces. It might be the reason why the pattern of context effect predicted by emotion seeds view was not found in the study carried out by Boucher (2014). So, in this chapter, I sought to examine whether the pattern of context effects proposed by emotion seeds views is applicable to six basic emotional facial expressions.

4.1 Experiment 1

In Experiment 1, bodily expressions were adopted as context stimuli to test whether the pattern of context effects from bodily expressions to recognition of facial expressions is consistent with the hypothesis of emotion seeds view.

4.1.1 Method

4.1.1.1 Participants

Twenty-seven new participants, ranging from 19 to 33 years, were recruited to take part in Experiment 1. Fifteen participants were female and 12 were male. All of the participants were right-handed. They all had given informed consent and were paid NZD 15 for participation. All had normal or corrected-to-normal vision.

4.1.1.2 Apparatus and Stimuli

The apparatus for this experiment is described in Chapter 2.

Stimuli of emotional faces were taken from the Karolinska Directed Emotional Faces set (KDEF) (Lundqvist et al., 1998). Two female and two male actors were selected, all expressing six basic emotions (Anger, Disgust, Fear, Happiness, Sadness, and Surprise). Images of bodily expressions were those evaluated in Experiment A. All stimuli of colour images of faces and bodies were converted into grey-scale images.

Face-body stimuli were edited using Adobe Photoshop CS5 (trial version). Three levels of perceptual similarity were created between the faces and body images: (a) Full similarity (congruent), a facial expression was placed on a body conveying the emotion that was the same as that was conveyed by the face (e.g. a disgust face on a disgust body); (b) High similarity, a face was placed on a body displaying the emotion that was highly similar to that displayed by the face (e.g. a disgust face on an anger body); (c) Low similarity, a face was placed on a body displaying the emotion less similar to that displayed the face (e.g. a disgust face on a fear body). The levels of perceptual similarity were based on the results in the study of Susskind et al. (2007) (details in Figure 1.4). The combinations of faces and bodies are listed in Table 4.1.

The stimuli, including face-body stimuli and isolated face images, were presented against a black background. The framed face-body stimuli were approximately 30 (width) x 19 (height) cm and the framed face images were approximately 7 (width) x 7 (height) cm. The size of isolated faces kept as the same as the size of faces in face-body stimuli. Isolated faces served as baseline stimuli and were presented in a separate block, with the same duration as face-body stimuli.

Table 4.1. The Combinations of Facial Expressions and Bodily Expressions

Facial Expressions	Levels of Perceptual Similarity (Bodily Expressions)		
	Full Similarity (Congruent)	High Similarity	Low Similarity
Anger	Anger	Disgust	Surprise
Disgust	Disgust	Anger	Fear
Fear	Fear	Surprise	Disgust
Happiness	Happiness	Surprise	Sadness
Sadness	Sadness	Fear	Happiness
Surprise	Surprise	Fear	Anger

4.1.1.3 Design and Procedure

The experiment used a 6 (facial expression: anger, disgust, fear, happiness, sadness, and surprise) x 4 (perceptual similarity: isolated face, full similarity, high similarity, and low similarity) within-subjects design.

The procedure for this experiment was the same as that described in Chapter 2. Participants were asked to click a mouse to choose one of six options to judge the expression shown on the faces when presented in the isolated block, and when in the face-body stimuli block they needed to judge the expression on faces while ignoring bodily expressions. In total, this experiment consisted of 960 trials, including 10 repetitions of 24 congruent trials (full similarity condition), 10 repetitions of 48 incongruent trials (24 trials of high similarity condition and 24 of low similarity condition), and 10 repetitions of 24 trials of isolated faces. Fourteen participants were shown the face-body stimuli block first, then the isolate face block. The rest participants were presented with the isolated face block first, then the face-body stimuli block. All participants were given a rest every 240 trials. Practice trials that were not used in the main test were given to participants to make them familiar with the experiment.

4.1.2 Results

4.1.2.1 Analysis for Accuracy of Facial Expressions in Bodily Context

Table 4.2 presents the mean percentage of accuracy of each facial expression paired with different levels of perceptual similarity bodies and isolated emotional faces. There were 30 trials excluded due to no response after a timeout (less than 0.12% of the trials). Accuracy was analysed with a 6 (facial expression: anger, disgust, fear, happiness, sadness, and surprise) x 4 (perceptual similarity: isolated face, full similarity, high similarity, and low similarity) repeated-measures ANOVA. *p* values were corrected by Greenhouse-Geisser epsilon, when sphericity was violated. Partial eta squared (η^2_p) is reported as effect size.

There was a main effect of facial expression, $F(3.11, 80.89) = 82.43$, $p < .001$, $\eta^2_p = .760$. Happy faces (97.56%) were recognised significantly better than all other emotional faces, all *p*s < .001. Surprise faces (75.00%) were significantly more accurate than sad faces (65.41%), $p = .016$, angry faces (60.94%), $p = .001$, fearful faces (40.65%), and disgusted faces (40.43%), both *ps* < .001. The accuracy of sad and angry faces was greater than that of fearful and disgusted faces, all *ps* < .001, and did not differ significantly from each other, $p = .157$. Fearful and disgusted faces were recognised the worst in all these facial expressions and did not find significant difference between each other ($p = .943$).

The main effect of perceptual similarity was also significant, $F(3, 78) = 296.60$, $p < .001$, $\eta^2_p = .919$. A facial expression under the condition of full similarity (congruent) (81.74%) was significantly more accurately recognised than under the condition of isolated face (67.63%), under the condition of low similarity (63.10%), and under the condition of high similarity (40.86%), all *ps* < .001. The accuracy for recognition of emotional faces was impaired the most under the condition of high similarity, compared to three other conditions, all *ps* < .001. The accuracy of emotional facial expression recognition under the condition of low similarity was significantly lower than that under the condition of isolated face, $p = .003$.

Due to significant interaction between facial expression and perceptual similarity $F(7.87, 204.56) = 30.16$, $p < .001$, $\eta^2_p = .537$, the analyses of accuracy of facial expression will be broken down into 6 subsections in each of which the results of six emotional faces will separately be reported. I conducted 6 one-way repeated-measures ANOVAs (one per each facial expression) with perceptual similarity as a within-subjects factor.

Table 4.2. Mean Percentages of Accuracy of Facial Expressions under Different Levels of Perceptual Similarity Bodies (%)

Facial Expression	Perceptual Similarity			
	Isolated Face	Full Similarity (Congruent)	High Similarity	Low Similarity
Anger	69.17(3.79)	87.64(1.89)	27.69(3.84)	59.25(3.30)
Disgust	33.74(2.69)	63.98(4.21)	20.73(3.08)	43.27(4.06)
Fear	41.30(4.42)	68.37(3.22)	21.82(3.42)	31.13(4.11)
Happiness	97.75(1.16)	96.76(2.06)	98.06(0.83)	97.69(1.29)
Sadness	85.00(2.58)	86.76(2.38)	21.64(2.65)	68.24(4.27)
Surprise	78.80(3.96)	86.94(2.38)	55.25(3.61)	79.00(3.59)

- SEM in parentheses

4.1.2.1a Analysis for Accurate Categorisation of Angry Facial Expressions

The accurate recognition of angry facial expressions was compared across the four levels of perceptual similarity to confirm whether a facial expression was facilitated by congruent bodily expressions, and to find out whether the recognition of facial expressions was impaired the most under the condition of high similarity. The results of one-way ANOVA revealed a main effect of perceptual similarity, $F(3, 78) = 89.08$, $p < .001$, $\eta^2_p = .774$. The accuracy of angry faces was significantly increased under the congruent bodily expressions (full similarity) (87.64%) compared to under the condition of isolated angry faces (69.17%), $p < .001$, and under the condition of high similarity (27.70%) and low similarity (59.25%), both *ps* < .001. The accuracy was significantly lower under the high similarity condition than under the isolated face condition, $p < .001$, and significantly lower under the low similarity condition than under the isolated face condition, $p = .018$. These two incongruent conditions (the high similarity condition and the low similarity condition) significantly differed from each other, $p < .001$. The pattern of results is consistent with the hypothesis proposed by emotion seeds view, with most impairment for recognition of angry facial expressions under the high similarity condition (disgusted bodily expressions).

4.1.2.1b Analysis for Accurate Categorisation of Disgusted Facial Expressions

A main effect of perceptual similarity for disgusted faces was found, $F(3, 78) = 45.98$, $p < .001$, $\eta^2_p = .639$. There was significant difference for accuracy between disgusted faces paired with congruent bodily expressions (63.98%) and disgusted faces presented separately (33.74%), $p < .001$. The accuracy of disgusted faces under the congruent condition was significantly larger than that of disgusted faces under the high similarity condition (20.73%) and under the low similarity condition (43.27%), both *ps* < .001. The accuracy of disgusted faces under high similarity condition was the lowest in these four levels of perceptual similarity, all *ps* < .001. The recognition of disgusted faces under the low similarity condition was significantly improved, compared to the condition of isolated disgusted faces ($p = .018$). The pattern of results is partially consistent with hypothesis of emotion seeds view. The worst recognition performance was observed for disgusted face paired with highly similar bodily expressions (angry bodies). However, the low similarity condition did not decrease the accuracy of disgusted faces, instead improved the accuracy of disgusted faces. The findings seem to be counter to intuition and are inconsistent with hypothesis of emotion seeds view which predicted that the condition of low similarity impairs the recognition of facial expressions, rather than improve it.

4.1.2.1c Analysis for Accurate Categorisation of Fearful Facial Expressions

There was a significant main effect of perceptual similarity for the accuracy of fearful facial expressions, $F(3, 78) = 81.54$, $p < .001$, $\eta^2_p = .759$. The accuracy of fearful facial expressions with congruent bodily expressions (68.37%) was significantly greater than isolated fearful faces (41.30%), fearful faces with highly similar bodily expressions (21.82%), and fearful faces with low similarity bodily expressions (31.13%), all *ps* < .001. The recognition of fearful faces under the high similarity condition was significantly worse than isolated fearful faces and under the low similarity condition, both *ps* < .001. Fearful faces paired with low similarity bodily expressions were recognised significantly worse than isolated fearful faces, $p = .006$. The pattern of results is consistent with emotion seeds view which predicted that the greatest interference effects for recognition of fearful faces occur under the condition of high similarity (surprise bodily expressions).

4.1.2.1d Analysis for Accurate Categorisation of Happy Facial Expressions

There was no main effect of perceptual similarity found for happy facial expressions, $F(1.47, 38.24) = 0.227$, $p = .729$. Happy facial expressions, overall, were recognised extremely high across all the four levels of perceptual similarity. The pattern of results is not consistent with emotion seeds view, with no interference effects under either of the two incongruent conditions.

4.1.2.1e Analysis for Accurate Categorisation of Sad Facial Expressions

There was a main effect of perceptual similarity for accurate recognition of sad facial expressions, $F(2.16, 56.09) = 165.59$, $p < .001$, $\eta^2_p = .864$. The accuracy of sad faces with congruent bodily expressions (86.76%) did not significantly differ from that of isolated sad faces (85.00%), $p = .371$. The accuracy of congruent face-body stimuli was significantly greater than the accuracy of sad faces with highly similar bodies (21.64%) and the accuracy of sad faces with low similarity bodies (68.24%), both *ps* < .001. Sad faces under the high similarity condition and under the low similarity were recognised significantly worse than sad faces presented in isolation, both *ps* < .001, and significant differences were observed between the high similarity condition and the low similarity condition, $p < .001$. The pattern of results is consistent with emotion seeds view which predicted that larger interference effects under the condition of high similarity than under the condition of low similarity.

4.1.2.1f Analysis for Accurate Categorisation of Surprised Facial Expressions

There was a main effect of perceptual similarity for accurate recognition of surprised facial expressions, $F(3, 78) = 41.54$, $p < .001$, $\eta^2_p = 0.615$. Surprised faces with congruent bodily expressions (86.94%) were recognised better than surprised faces presented in isolation (78.80%), $p = .004$, surprised faces with fearful bodily expressions (the high similarity condition) (55.25%), $p < .001$, and surprised faces with angry bodily expressions (the low similarity condition (79.00%), $p = .003$. The lowest accuracy was found when surprised faces were paired with fearful bodily expressions (the high similarity condition), which was significantly lower than the conditions of isolated surprised faces and surprised faces paired with angry bodily expressions (low similarity condition), both $ps < .001$. There was no significant difference between surprised faces paired with angry bodily expressions (the low similarity condition) and surprised faces in isolation, $p = .946$. The pattern of results is consistent with hypothesis by emotion seeds view, with greater interference effects elicited from high similarity bodies than low similarity bodies on recognition of surprised facial expressions.

4.1.2.2 Errors Analyses

According to emotion seeds view, shared perceptual similarity between facial expressions is called emotion seeds. It is predicted that the more emotion seeds shared between an emotion on a face and an emotion on an accompanied context, the greater probability to mis-categorise the facial expression as the emotion conveyed by the concurrently presented context. In other words, a facial expression under the condition of highly similar context would be more likely to be mis-categorised as the emotion conveyed by the context than any other incongruent emotions. For example, the percentage of mis-categorising an angry face paired with a disgusted body (that is, under the high similarity context) as disgusted would be larger than the percentage of mis-categorising the angry face as surprised, fearful, happy, or sad.

Emotion seeds view proposed that emotion seeds remain latent when emotional facial expressions are presented in isolation. And emotion seeds exert influence on recognition of facial expressions only when emotional faces are paired with appropriate context. In other words, the percentage of mis-categorisation of a facial expression paired with high similar context as the emotion conveyed by the context would be greater, compared to the condition of low similarity and the condition of isolated faces.

Additionally, based on emotion seeds view, the percentage of mis-categorisation of a facial expression as the emotion delivered by the concurrently presented context will be reduced as the level of perceptual similarity decreases. That is, a facial expression paired with a high similarity context will be more likely to be mis-categorised as the emotion conveyed by the context than the same facial expression paired with a low similarity context. For example, the percentage of mis-categorising an angry face paired with disgusted bodies as disgust will be larger than the percentage of mis-categorising the angry face paired with surprised bodies as surprise. Target mis-categorisation refers to the mis-categorisation of a facial expression as the emotion conveyed by the concurrently presented context. Table 4.3 lists the target mis-categorisation of each facial expression paired with high similarity context and low similarity context.

Table 4.3. Target Mis-categorisation of Each Facial Expression under High Similarity Contexts and Low Similarity Contexts

	Target Mis-categorisation	
Facial Expression	High Similarity Context	Low Similarity Context
Anger	Disgust	Surprise
Disgust	Anger	Fear
Fear	Surprise	Disgust
Happiness	Surprise	Sadness
Sadness	Fear	Happiness
Surprise	Fear	Anger

Error analyses were performed for each of the six facial expressions in the following subsections.

4.1.2.2a Analysis for Mis-categorisation of Angry Facial Expressions

The percentage of mis-categorisation of angry facial expressions paired with disgusted bodies is listed in Table 4.4. The one-way repeated-measures ANOVA (emotion: disgust, fear, happiness, sadness, or surprise) was conducted on the mis-categorisation of angry facial expressions under the condition of high similarity. There was a main effect of emotion on the mis-categorisation of angry faces, $F(2.03, 52.73) = 36.03$, $p < .001$, $\eta^2_p = .581$. Pairwise comparisons (Fisher's LSD) revealed that the mis-categorisation of angry facial expressions as disgusted (38.95%) was significantly higher than any other incongruent emotions, all *ps* < .001. The mis-categorisation of angry facial expressions as happy (0%) was significantly lower than as fearful (10.29%), as sad (13.54%), and as surprised (9.45%), all *ps* < .001. There were no other significant differences between each pair of the mis-categorisation of angry faces as fearful, as sad, and as surprised, all *ps* > .10. The results are consistent with emotion seeds view which predicted that a facial expression paired with highly similar context is more likely to be mis-categorised as the emotion conveyed by highly similar context than as any other incongruent emotions.

Table 4.4. Mean Percentage of Mis-categorisation of Facial Expressions Paired with High Similarity Bodies (%)

Face-body Composite	Mis-categorisation					
	Anger	Disgust	Fear	Happiness	Sadness	Surprise
Anger-Disgust		38.95 (3.83)	10.29 (1.95)	0.00	13.54 (2.34)	9.45 (1.62)
Disgust-Anger	21.66 (2.21)		2.73 (0.92)	8.82 (2.04)	44.37 (2.92)	1.69 (0.77)
Fear-Surprise	1.58 (0.67)	1.67 (0.48)		0.37 (0.17)	26.68 (2.44)	47.89 (3.12)
Sadness-Fear	3.54 (0.94)	13.82 (3.10)	56.93 (3.85)	0.19 (0.13)		3.80 (1.00)
Surprise-Fear	0.74 (0.42)	0.56 (0.24)	41.97 (3.48)	0.56 (0.24)	0.84 (0.50)	

- SEM in parentheses

Table 4.5 presents the percentage of mis-categorisation of angry faces as disgust across the conditions of perceptual similarity. A one-way repeated-measures ANOVA (perceptual similarity: isolated face, high similarity, low similarity) was conducted for the mis-categorisation of angry faces as disgust. A main effect of perceptual similarity was significant, $F(1.49, 38.69) = 54.35$, $p < .001$, $\eta^2_p = .676$. Pairwise comparisons (Fisher's LSD) showed that the percentage of mis-categorisation of angry faces paired with disgusted bodies (high similarity context) as disgust (38.95%) was significantly more than that of mis-categorisation of angry faces paired with surprised bodies (low similarity context) as disgust (18.06%) and that of mis-categorisation of isolated angry faces as disgust (12.41%), both *ps* $< .001$. And the percentage of mis-categorisation of angry faces as disgust was significantly greater under the condition of low similarity (that is, surprised bodies) than that for isolated angry faces, $p = .003$. The results confirm that emotion seeds exert greater influence on recognition of angry facial expressions when paired with disgusted bodies (that is, under the condition of high similarity context).

Table 4.5. Mean Percentage of Mis-categorisation of Facial Expressions as the Emotion by High Similarity Bodies across the Conditions of Perceptual Similarity (%)

Face-Miscategorisation	Perceptual Similarity		
	Isolated Face	High Similarity	Low Similarity
Anger-Disgust	12.41 (1.69)	38.95 (3.83)	18.06 (2.63)
Disgust-Anger	5.65 (1.17)	21.66 (2.21)	1.57 (0.60)
Fear-Surprise	37.50 (3.15)	47.89 (3.12)	46.81 (3.76)
Sadness-Fear	2.13 (0.73)	56.93 (3.85)	1.85 (0.69)
Surprise-Fear	12.50 (3.15)	41.97 (3.48)	6.13 (1.56)

- SEM in Parentheses

The separate paired-sample t-test for anger trials confirm that the percentage of mis-categorisation as disgusted when angry facial expressions were paired with disgusted bodies (38.95%) was significantly greater than that of mis-categorisation as surprised when angry facial expressions were paired with surprised bodies (1.30%), $t(26) = 9.69$, $p < .001$, $d = 1.87$. The results are consistent with emotion seeds view, with greater percentage of target mis-categorisation elicited by high similarity bodies than by low similarity bodies. In other words, the percentage of target mis-categorisation is reduced as the level of perceptual similarity is decreased.

4.1.2.2b Analysis for Mis-categorisation of Disgusted Facial Expressions

Table 4.4 presents the percentage of mis-categorisations of disgusted facial expressions paired with angry bodies. A one-way repeated-measures ANOVA (emotion: anger, fear, happiness, sadness, or surprise) was performed on the mis-categorisation of disgusted facial expressions under the condition of high similarity. There was a main effect of emotion on the mis-categorisation of disgusted faces, $F(2.42, 62.93) = 74.20$, $p < .001$, $\eta^2_p = .741$. Pairwise comparisons (Fisher's LSD) revealed that the mis-categorisation of disgusted faces as anger (21.66%) was significantly higher than the mis-categorisation of disgusted faces as fearful (2.73%), happy (8.82%), and surprised (1.69%), all *ps* < .001, however was significantly lower than the mis-categorisation as sad (44.37%), $p < .001$. The percentage of mis-categorisation of disgusted faces as sad was significantly higher than that of mis-categorisation of disgusted faces as fearful, happy, and surprised, all *ps* $< .001$. There was significant difference between the mis-categorisation of disgusted faces as fearful and as happy, $p = .017$, and no significant difference between the mis-categorisation as fearful and as surprised, $p = .314$. The percentage of mis-categorisation as happy was significant different from that as surprised, $p = .004$. The results are inconsistent with emotion seeds view which predicted that a facial expression paired with highly similar context is more likely to be mis-categorised as the emotion conveyed by highly similar context than as any other incongruent emotions.

The percentage of mis-categorisation of disgust faces as anger across the conditions of perceptual similarity is listed in Table 4.5. A one-way repeated-measures ANOVA (perceptual similarity: isolated face, high similarity, low similarity) was conducted for the mis-categorisation of disgusted faces as anger. A main effect of perceptual similarity was significant, $F(1.59, 41.42) = 64.64$, $p < .001$, $\eta^2_p = .713$. Pairwise comparisons (Fisher's LSD) showed that the percentage of mis-categorisation of disgusted faces paired with angry bodies (high similarity context) as anger (21.66%) was significantly more than that of mis-categorisation of disgusted faces paired with fearful bodies (low similarity context) as anger (5.65%) and that of mis-categorisation of isolated disgusted faces as anger (1.57%), both *ps* $< .001$. And the percentage of mis-categorisation of disgusted faces as angry was significantly greater under the condition of low similarity than that for isolated disgusted faces, $p = .005$. The results confirm that emotion seeds exert greater influence on recognition of disgusted faces when paired with angry bodies (that is, in the condition of high similarity context).

The paired-sample t-test found that the percentage of mis-categorisation of disgusted facial expressions with fearful bodies as fearful (39.31%) was significantly greater than that of mis-categorisation of disgusted faces with angry bodies as angry (21.66%), $t(26) = 3.708$, $p = .001$, $d = 0.72$. The results are inconsistent with emotion seeds view, with larger percentage of target mis-categorisation elicited by low similarity bodies than by high similarity bodies.

4.1.2.2c Analysis for Mis-categorisation of Fearful Facial Expressions

Table 4.4 presents the percentage of mis-categorisations of fearful facial expressions paired with surprised bodies. A one-way repeated-measures ANOVA (emotion: anger, disgust, happiness, sadness, or surprise) was conducted on the mis-categorisation of fearful facial expressions under the condition of high similarity. There was a main effect of emotion on the mis-categorisation of fearful faces, $F(1.68, 43.72) = 126.99$, $p < .001$, $\eta^2_p = .830$. Pairwise comparisons (Fisher's LSD) showed that the mis-categorisation of fearful faces as surprised (47.89%) was significantly higher than any other incongruent emotions, all *ps* < .001. The mis-categorisation of fearful facial expressions as sad (26.68%) was significantly greater than as angry (1.58%), disgusted (1.67%), and happy (0.37%), all *ps* < .001. There was significant difference between the mis-categorisation of fearful faces as disgusted and as happy, p = .020. The percentage of mis-categorisation of fearful facial expressions as angry was marginally higher that of mis-categorisation of fearful facial expressions as happy, p = .067. There was no significant difference between the percentage of mis-categorisation fearful facial expressions as angry and that of mis-categorisation fearful facial expressions as disgusted, p = .913. The results are consistent with emotion seeds view which predicted that a facial expression paired with highly similar context is more likely to be mis-categorised as the emotion conveyed by highly similar context than as any other incongruent emotions.

The percentage of mis-categorisation of fearful faces as surprise across the conditions of perceptual similarity is listed in Table 4.5. A one-way repeated-measures ANOVA (perceptual similarity: isolated face, high similarity, low similarity) was conducted for the mis-categorisation of fearful faces as surprise. A main effect of perceptual similarity was significant, $F(2, 52) = 10.83$, $p < .001$, $\eta^2_p = .294$. Pairwise comparisons (Fisher's LSD) showed that the percentage of mis-categorisation of fearful faces paired with surprised bodies (high similarity condition) as surprise (47.89%) was significantly more than that of mis-categorisation of isolated fearful faces as surprise (37.50%), $p < .001$, but was not significantly different from that of mis-categorisation of fearful faces paired with disgusted bodies (low similarity condition) as surprise (46.81%), p = .622. The percentage of mis-categorisation of fearful faces as surprised was significantly greater under the condition of low similarity than that for isolated fearful faces, p = .002. The results that no significant differences between the conditions of high similarity and low similarity are not consistent with emotion seeds view which predicted greater influence on recognition of fearful faces will occur when paired with surprised bodies (that is, in the condition of high similar context).

The paired-sample t-test for fear trials showed the percentage of mis-categorisation of fearful faces paired with surprised bodies as surprised (47.89%) was significantly greater than that of mis-categorisation of fearful faces paired with disgusted bodies as disgusted (5.74%), $t(26) = 12.62$, $p < .001$, $d = 2.42$. The results are consistent with emotion seeds view, with greater percentage of target mis-categorisation elicited by high similarity context than by low similarity context.

4.1.2.2d Analysis for Mis-categorisation of Happy Facial Expressions

The analysis for accuracy revealed that the recognition of happy facial expressions was not significantly impaired by incongruent bodies. Therefore, the error analysis was not performed on happy facial expressions.

4.1.2.2e Analysis for Mis-categorisation of Sad Facial Expressions

Table 4.4 presents the percentage of mis-categorisations of sad facial expressions paired with fearful bodies. A one-way repeated-measures ANOVA (emotion: anger, disgust, fear, happiness, or surprise) was conducted on the mis-categorisation of sad facial expressions under the condition of high similarity. There was a main effect of emotion on the mis-categorisation of sad faces, $F(1.43, 37.05) = 89.84$, $p < .001$, $\eta^2_p = .776$. Pairwise comparisons (Fisher's LSD) revealed that the mis-categorisation of sad faces as fearful (56.93%) was significantly higher than any other incongruent emotions, all *ps* < .001. The mis-categorisation of sad facial expressions as disgusted (13.82%) was significantly more than that as angry (3.54%), $p = .004$, as happy (0.19%), $p < .001$, and as surprised (3.80%), $p = .007$. The mis-categorisation of sad facial expressions as angry was significantly more than that as happy, $p = .001$. And the mis-categorisation of sad facial expressions as surprised was significantly more than that as happy, $p = .002$. There was no significant difference between the mis-categorisation as angry and as surprised, $p = .858$. The results are consistent with emotion seeds view which predicted that a facial expression paired with highly similarity context is more likely to be mis-categorised as the emotion conveyed by highly similar context than as any other incongruent emotions.

The percentage of mis-categorisation of sad faces as fear across the conditions of perceptual similarity is listed in Table 4.5. A one-way repeated-measures ANOVA (perceptual similarity: isolated face, high similarity, low similarity) was conducted for the mis-categorisation of sad faces as fear. A main effect of perceptual similarity was significant, $F(1.07, 27.78) = 207.23$, $p < .001$, $\eta^2_p = .889$. Pairwise comparisons (Fisher's LSD) showed that the percentage of mis-categorisation of sad faces paired with fearful bodies (high similarity context) as fear (56.93%) was significantly more than that of mis-categorisation of sad faces paired with happy bodies (low similarity context) as fear (1.85%) and that of mis-categorisation of isolated sad faces as fear (2.13%), both *ps* < .001. The percentage of mis-categorisation of sad faces as fear did not significantly differ between the condition of low similarity and isolated sad faces, *p* = .731. The results confirm that emotion seeds exert greater influence on recognition of sad faces when paired with fearful bodies (that is, in the condition of high similarity context).

The paired-sample t-test for sadness trials showed that the mis-categorisation of sad faces paired with fearful bodies as fearful (56.93%) was significantly greater than the mis-categorisation of sad faces paired with happy bodies as happy (8.27%), $t(26) = 11.84$, $p < .001$, $d = 2.28$. The results are consistent with emotion seeds view, with greater percentage of target mis-categorisation caused by high similarity context than by low similarity context.

4.1.2.2f Analysis for Mis-categorisation of Surprised Facial Expressions

Table 4.4 presents the percentage of mis-categorisations of surprised facial expressions paired with fearful bodies. A one-way repeated-measures ANOVA (emotion: anger, disgust, fear, happiness, or sad) was conducted on the mis-categorisation of surprised facial expressions under the condition of high similarity.

There was a main effect of emotion on the mis-categorisation of surprised faces, $F(1.09, 28.25) = 135.75$, $p < .001$, $\eta^2_p = .839$. Pairwise comparisons (Fisher's LSD) revealed that the mis-categorisation of surprised faces as fearful (41.97%) was significantly more than any other incongruent emotions, all *ps* < .001. There were no other significant differences between each pair of the other four emotions (angry, 0.74%, disgusted, 0.56%, happy, 0.56%, and sad, 0.84%), all *ps* > .10. The results are consistent with emotion seeds view which predicted that a facial expression paired with highly similar context is more likely to be mis-categorised as the emotion conveyed by highly similar context than as any other incongruent emotions.

The percentage of mis-categorisation of surprised faces as fear across the conditions of perceptual similarity is listed in Table 4.5. A one-way repeated-measures ANOVA (perceptual similarity: isolated face, high similarity, low similarity) was conducted for the mis-categorisation of surprised faces as fear. A main effect of perceptual similarity was significant, $F(2, 52) = 93.01$, $p < .001$, $\eta^2_p = 0.782$. Pairwise comparisons (Fisher's LSD) showed that the percentage of mis-categorisation of surprised faces paired with fearful bodies (high similar context) as fear (41.97%) was significantly more than that of mis-categorisation of surprised faces paired with angry bodies (low similar context) as fear (6.13%) and that of mis-categorisation of isolated sad faces as fear (12.50%), both *ps* < .001. The percentage of mis-categorisation of isolated surprised faces as fear was significantly greater than that of mis-categorisation of surprised faces paired with angry bodies as fear, $p = .017$. The results confirm that emotion seeds exert greater influence on recognition of surprised faces when paired with fearful bodies (that is, in the condition of high similar context).

The paired-sample t-test for surprise trials revealed that the percentage of mis-categorisation of surprised faces paired with fearful bodies as fearful (41.97%) was significantly greater than that of mis-categorisation of surprised faces paired with angry bodies as angry (8.37%), $t(26) = 9.96$, $p < .001$, $d = 1.92$. The results are consistent with emotion seeds view, with greater percentage of target mis-categorisation caused by high similar context than by low similar context.

4.1.3 Discussion

In all the six facial expressions, happy faces were recognised most accurately. It is consistent with previous studies which found happy facial expressions are the easiest facial expressions to be recognised and can be recognised more accurately and faster than any other emotional facial expressions (Ambadar, Schooler, & Cohn, 2005; Kirita & Endo, 1995). The accuracy rates of disgusted and fearful faces were the lowest, which is consistent with the results in other studies (Biehl et al., 1997; Blair et al., 2004; Castelli, 2005; Martinez & Du, 2012; Russell, 1994). These studies indicated that although disgusted and fearful facial expressions were the most difficult to recognise, it is important for survival to recognise fearful and disgusted facial expressions. Compared to four other facial expressions, these two facial expressions can convey dangerous signals from environment. Fearful facial expressions can alert conspecifics potential threat, in order to keep away from danger (Marsh, Ambady, & Kleck, 2005; Öhman & Mineka, 2001; Shariff & Tracy, 2011). Disgusted facial expressions act as signals of warning organism about toxic or contaminated food, in order to prevent organism from eating it to avoid disease or danger (Chapman, Kim, Susskind, & Anderson, 2009; Rozin, Lowery, & Ebert, 1994). It is of advantage to recognise fear and disgust far away from us. Thus, it might rely on distal cues (e.g. bodily expressions, vocal expressions) more than on facial expressions to recognise fear and disgusted.

The analyses for accuracy rates showed that the recognition of angry, disgusted, fearful, surprised facial expressions paired with congruent bodily expressions was increased, compared to isolated angry, disgusted, fearful, surprised facial expressions. The results are consistent with the results in previous studies (Aviezer et al., 2008b; Meeren et al., 2005; van den Stock et al., 2007; Mondloch, 2012; Mondloch et al., 2013b; Mondloch, Horner, & Mian, 2013a; van de Riet & de Gelder, 2008). The facilitation effects were not found for happy and sad faces, which replicated the results in the study by Boucher (2014). In the study by Boucher (2014), the facilitation effects were only observed when fearful faces paired with fearful bodies, but were not observed for the other five facial expressions paired with congruent bodies.

The following discussion excludes happy facial expressions, because the accuracy and error analyses revealed that there were neither facilitation effects from congruent bodies nor interference effects from incongruent bodies. For the other five facial expressions, the accuracy analyses revealed that each facial expression when paired with a high similarity body was recognised the worst, compared to the same facial expression, a facial expression with a congruent body, and a facial expression with a low similarity body. The pattern of results is consistent with emotion seeds view, which predicted that maximal influence under the high similarity condition, and minimal influence under the low similarity condition.

The three error analyses have been performed to explore: (a) whether a facial expression paired with a high similarity context is more likely to be mis-categorised as the emotion conveyed by high similarity context than any other incongruent emotions; (b) whether emotion seeds exert greater influence on recognition of a facial expression under the condition of high similarity context, compared to the same facial expression in isolation and under the condition of low similarity context; (c) whether the percentage of mis-categorisation of a facial expression as the emotion by concurrently presented context is reduced as the level of perceptual similarity decreases (that is, the percentage of target mis-categorisation is greater under the condition of high similarity context than under the condition of low similarity context.

The error analysis carried on the emotional faces paired with high similarity bodies showed that except disgusted facial expression, four other facial expressions (anger, fear, sadness, and surprise) were mis-categorised more as the emotions conveyed by high similarity bodies than any other incongruent emotions. The pattern of the results is consistent with emotion seeds view. Under the condition of high similarity bodies, facial expressions are more likely to be mis-categorised as the emotions conveyed by high similarity contexts.

A second error analysis was performed on mis-categorisation of facial expressions as the emotions by high similarity bodies across the conditions of high similar context, low similar context, and isolated faces. The results showed that all the five facial expressions (anger, disgust, fear, sadness, and surprise) paired with high similarity contexts were mis-categorised more as the emotions conveyed by high similarity contexts, compared with the same five facial expressions in isolation. For the comparison between facial expressions with high similarity context and low similarity context, the percentage of mis-categorisation of four facial expressions (anger, disgust, sadness, and surprise) as the emotions conveyed by high similarity contexts was significantly greater under the condition of high similarity context than under the condition of low similarity context. However, there was no significant difference of mis-categorising fearful faces as surprise between high similarity context and low similarity context. The results showed that except fearful facial expressions, emotion seeds have greater impact on the recognition of four other facial expressions (anger, disgust, sadness, and surprise) when they are presented with the high similarity contexts.

Moreover, the last error analysis showed that except disgusted faces, the percentage of target mis-categorisation of the other four emotional faces (anger, fear, sadness, and surprise) paired with high similarity contexts was greater than that of target mis-categorisation of emotional faces paired with low similarity contexts. For example, the percentage of target mis-categorisation (that is, disgust) of angry faces paired with disgusted bodies was greater than that of target mis-categorisation (that is, surprise) of angry faces paired with surprised bodies. The target mis-categorisation of each facial expression lists above in Table 4.3. The findings are partially consistent with emotion seeds view which predicted that the percentage of mis-categorisation of facial expressions as the emotions by the concurrently presented context was reduced as the levels of perceptual similarity were decreased.

It is interesting and noteworthy that the accuracy of disgusted faces was significantly increased when disgusted faces are paired with fearful bodies (low similarity context), compared to isolated disgusted faces. Another interesting result is that the mis-categorisation of disgusted facial expressions paired with fearful bodies (low similarity condition) as fear was significantly greater than that of disgusted facial expressions paired with angry bodies (high similarity condition) as anger. Emotion seeds view fails to explain the unpredicted results.

In sum, facilitation effects from congruent bodies was found for angry, fearful, disgusted, and surprised facial expressions, but was not observed for happy and sad facial expressions. The pattern of the accuracy analyses for emotional faces with incongruent bodily expressions, except happy facial expression, is consistent with emotion seeds view which predicted that interference effects from high similarity bodies on the recognition of facial expressions were greater than interference effects from low similarity bodies on the recognition of facial expressions. Further, the error analyses showed that the pattern of interference effects under the incongruent conditions (that is, high similarity condition and low similarity condition) predicted by emotion seeds view was fully supported by the results of angry, surprised, and sad facial expressions, whereas the predictions of emotion seeds view were partially supported by the results of disgusted and fearful facial expressions.

4.2 Experiment 2

Experiment 2 adopted vocal expressions as context stimuli. The aim of this experiment was to examine whether the pattern of context effects from vocal expressions on recognition of facial expressions is consistent with the hypothesis of emotion seeds view.

4.2.1 Method

4.2.1.1 Participants

Twenty-four new participants, ranging from 19 to 40 years, were recruited to take part in Experiment 2. Twelve participants were female and 12 were male. All of the participants were right-handed. They all had given informed consent and were paid NZD 15 for participation. All reported normal or corrected-to-normal vision and normal hearing.

4.2.1.2 Apparatus and Stimuli

This experiment was run on the same computer as that described in chapter 2. The auditory stimuli were presented via the same headphone as in Experiment B at self-adjusted comfortable level.

Face stimuli were the same as those used in Experiment 1. Sound stimuli (1 female and 1 male) were from MAV set (Belin et al., 2008), which were evaluated in experiment B. Face-sound stimuli created three levels of perceptual similarity, which is the same as Experiment 1. Isolated faces, the same as those in face-sound stimuli, served as baseline stimuli and were presented in a separate block.

4.2.1.3 Design and Procedure

The experiment used a 6 x 4 within-subjects design, with the factor of facial expression (anger, disgust, fear, happiness, sadness, and surprise) and the factor of perceptual similarity (isolated face, full similarity, high similarity, and low similarity).

The procedure for this experiment was almost identical to that in Experiment 1. In face-sound stimuli block, facial expression images and affect bursts were presented simultaneously, with the same duration of 1830ms. Short vocal stimulus was played in a loop as that in Experiment B. In isolated face block, face images were presented for 1830ms. Participants were asked to use the mouse to click on one of six options (anger, disgust, fear, happiness, sadness, and surprise) to judge expressions shown on the faces when presented in an isolation block, and when in a face-sound stimuli block they needed to judge expressions on faces while ignoring accompanied sounds. In total, this experiment consisted of 960 trials, including 10 repetitions of 24 congruent trials (full similarity condition), 10 repetitions of 48 incongruent trials (24 trials of high similarity condition and 24 of low similarity condition), and 10 repetitions of 24 trials of isolated faces. Half participants (12 participants) were presented with isolated face block first, and then face-sound stimuli block; the other half participants (12 participants) were presented with face-sound stimuli block first, and then isolated face block. All participants were given a rest every 240 trials. Practice trials were given to participants to make them familiar with the experiment.

4.2.2 Results

4.2.2.1 Analysis for Accuracy of Facial Expressions in Auditory Context

Table 4.6 presents the percentage of accuracy of each facial expression paired with different levels of perceptual similar sounds and isolated emotional faces. There were 22 trials excluded due to no response after a timeout (less than 0.10% of the trials). Accuracy was analysed with a 6 (facial expression: anger, disgust, fear, happiness, sadness, and surprise) x 4 (perceptual similarity: isolated face, full similarity, high similarity, and low similarity) repeated-measures ANOVA. *p* value were corrected by Greenhouse-Geisser epsilon, when sphericity was violated. Partial eta squared (η^2_p) is reported as effect size.

There was a main effect of facial expression, $F(3.54, 81.32) = 25.48$, $p < .001$, $\eta^2_p = .526$. Happy faces (86.43%) were recognised significantly better than all the other five emotional faces, all *p*s < .001, which is the same as the results in Experiment 1. The percentage of accuracy of sad faces (77.35%) were significantly greater than the percentage of accuracy of angry faces (67.57%), $p = .003$, surprised faces (62.22%), $p = .001$, disgusted faces (49.16%), and fearful faces (48.28%), both *ps* < .001. The accuracy of angry faces was significantly greater than that of disgusted faces, $p = .001$, and fearful faces, $p < .001$, but was not significant different from surprised faces, $p = .215$. The recognition of surprised faces was significantly better than disgusted faces,

p = .010, and fearful faces, p = .019. Disgusted and fearful faces were recognised the worst of all these facial expressions and there were no significant differences between each other (p = .878).

A main effect of perceptual similarity was also significant, $F(1.14, 26.19) = 28.67$, $p < .001$, $\eta^2_p = .555$. A facial expression under the condition of full similarity (81.52%) was recognised significantly more accurately than under the condition of isolated face (72.49%), under the low similarity condition (54.18%), and under the high similarity condition (52.47%), all *ps* < .001. The accuracy of facial expression recognition under the conditions of low similarity and high similarity was significantly less than that under the condition of isolated face, both *ps* < .001, and did not significantly differ from each other, p = .149.

Due to significant interaction between facial expression and perceptual similarity $F(5.36, 123.31) = 11.04$, $p < .001$, $\eta^2_p = .324$, the analysis of accuracy of facial expressions will also be broken down into 6 subsections in each of which the results of six emotional faces will separately be reported. I conducted 6 one-way repeated-measures ANOVAs (one per each facial expression) with perceptual similarity as within-subjects factor. The accurate rates for each facial expression were compared across the four levels of perceptual similarity to confirm whether a facial expression was facilitated by congruent vocal expressions, and to find out whether the recognition of facial expressions was impaired the most under the condition of high similarity.

Table 4.6. Mean Percentages of Accuracy of Facial Expressions under Different Levels of Perceptual Similarity Sound (%)

Facial Expression	Perceptual Similarity			
	Isolated Face	Full Similarity (Congruent)	High Similarity	Low Similarity
Anger	76.46 (2.65)	83.41 (2.42)	47.59 (5.72)	62.81 (4.24)
Disgust	44.09 (5.37)	81.77 (4.70)	38.26 (5.98)	32.50 (5.86)
Fear	51.04 (5.03)	64.38 (4.16)	50.23 (4.25)	27.46 (4.64)
Happiness	98.44 (0.58)	97.60 (1.15)	68.96 (7.87)	80.73 (6.03)
Sadness	89.58 (2.29)	91.77 (1.87)	57.20 (6.99)	70.83 (5.48)
Surprise	75.31 (3.75)	70.21 (4.25)	52.59 (5.49)	50.76 (5.96)

- SEM in Parentheses.

4.2.2.1a Analysis for Accurate Categorisation of Angry Facial Expressions

A one-way repeated measures ANOVA revealed a main effect of perceptual similarity for angry faces, $F(1.67, 38.44) = 27.17$, $p < .001$, $\eta^2_p = .542$. The accuracy of angry facial expressions paired with angry sounds (full similarity, 83.41%) than that of isolated angry facial expressions (74.46%), $p = .018$, that of angry facial expressions paired with surprised sounds (low similarity, 62.81%), and that of angry facial expressions paired with disgusted sounds (high similarity, 47.59%), both *ps* $< .001$. The accuracy of angry facial expressions paired with high similarity context was significantly less than that of isolated angry faces, $p < .001$. The accuracy of angry facial expressions paired with low similarity context was significantly less than that of isolated angry faces, $p = .002$. These two incongruent conditions significantly differed from each other, $p < .001$. The pattern of results is consistent with the hypothesis proposed by emotion seeds view, with most impairment for recognition of angry facial expressions paired with disgusted sounds (that is, under the condition of high similarity).

4.2.2.1b Analysis for Accurate Categorisation of disgusted Facial Expressions

A main effect of perceptual similarity for disgusted faces was found, $F(1.97, 45.37) = 35.28$, $p < .001$, $\eta^2_p = .605$. There was significant difference between the accuracy of disgusted faces paired with disgusted sound (81.77%) and the accuracy of disgusted faces presented in isolation (44.09%), $p < .001$. The accuracy of disgusted faces under the congruent condition was significantly larger than the accuracy of disgusted faces under both incongruent conditions (the high similarity condition, 38.26% and the low similarity condition, 32.50%), both *ps* $< .001$. The accuracy of disgusted faces under the low similarity condition (fearful sounds) were significantly lower than that of isolated disgusted faces, $p = .010$, and was marginally different from disgusted faces under the high similarity condition (angry sounds), *p* = .059. There was no significant difference between disgusted faces in isolation and disgusted faces accompanied with angry sounds (high similarity), *p* = . 218. The pattern of the results is inconsistent with hypothesis of emotion seeds view which predicted that larger interference effects elicited by angry sounds (high similarity context) than by fearful sounds (low similarity context) when recognising disgusted facial expressions.

4.2.2.1c Analysis for Accurate Categorisation of Fearful Facial Expressions

There was a main effect of perceptual similarity for the accuracy of fearful facial expressions, $F(1.97, 45.20) = 21.41$, $p < .001$, $\eta^2_p = 0.482$. The accuracy of fearful facial expressions paired with fearful sounds (64.38%) was significantly greater than isolated fearful facial expressions (51.04%), *p* = .001, greater than that of fearful facial expressions paired with surprised sounds (high similarity condition, 50.23%), *p* = .002, and greater than that of fearful facial expressions paired with disgusted sounds (low similarity condition, 27.46%), $p < .001$. The recognition of fearful faces paired with disgusted sounds was significantly lower than all the other conditions, all *ps* < .001. There was no significant difference between isolated fearful faces and fearful faces paired with surprised sounds, *p* = .838. The pattern of the results is inconsistent with emotion seeds view, without significant difference between interference effects elicited by high similarity context and those elicited by low similarity context.

4.2.2.1d Analysis for Accurate Categorisation of Happy Facial Expressions

There was a main effect of perceptual similarity found for happy facial expressions, $F(1.38, 31.79) = 11.71$, $p = .001$, $\eta^2_p = .337$. The accuracy of happy faces with happy sound (97.60%) was not significantly increased compared to isolate happy faces (98.44%), $p = .477$. The accuracy of happy facial expressions paired with happy sounds was significantly larger than that of happy facial expressions paired with surprised sounds (high similar context, 68.96%), $p = .001$, and larger than that of happy facial expressions paired with sad sounds (low similar context, 80.73%), $p = .007$. The accuracy of happy facial expressions paired with surprised sounds was significantly less than that of isolated happy facial expressions, $p = .001$, and was significantly less than that of happy facial expressions paired with sad sounds, $p = .007$. There were significant differences of the accuracy between happy facial expressions paired with surprised sounds and those paired with sad sounds, $p = .014$. The pattern of results is consistent with emotion seeds view, with greater interference effects under high similarity condition than under low similarity condition.

4.2.2.1e Analysis for Accurate Categorisation of Sad Facial Expressions

There was a main effect of perceptual similarity for the accurate recognition of sad facial expressions, $F(1.90, 43.70) = 17.53$, $p < .001$, $\eta^2_p = .433$. The accuracy of sad faces with sad sounds (91.77%) did not significantly differ from that of isolated sad faces (89.58%), p = .303. The accuracy of sad facial expressions paired with sad sounds was significantly larger than that of sad facial expressions paired with fearful sounds (high similarity context, 57.20%) and that of sad facial expressions paired with happy sounds (low similarity context, 70.83%), both ps < .001. The accuracy of sad facial expressions paired with fearful sounds was significantly less than that of isolated sad facial expressions, p < .001, and was significantly less than that of sad facial expressions paired with happy sounds, p = .030. The accuracy of sad facial expressions paired with happy sounds was significantly less than that of isolated sad facial expressions, p = .001. The pattern of the results is consistent with emotion seeds view which predicted that larger interference effects elicited by fearful sounds (high similarity context) than by happy sounds (low similarity context).

4.2.2.1f Analysis for Accurate Categorisation of Surprised Facial Expressions

There was a main effect of perceptual similarity for the accurate recognition of surprised facial expressions, $F(1.49, 34.20) = 13.50$, $p < .001$, $\eta^2_p = .370$. The accuracy of isolated surprised facial expressions (75.31%) was marginally larger than that of surprised facial expressions paired with surprised sounds (70.21%), p = .091. The accuracy of surprised facial expressions paired with surprised sounds was significantly larger than that of surprised facial expressions paired with fearful sounds (high similarity, 52.59%), p = .002, was also significantly larger than that of surprised facial expressions paired with angry sounds (low similarity, 50.76%), p = .001. The accuracy of isolated surprised facial expressions was significantly larger than that of surprised facial expressions paired with fearful sounds, $p = .001$, and larger than that of surprised facial expressions paired with angry sounds, $p < .001$. The accuracy of surprised facial expressions paired with fearful sounds was not significantly different from that of surprised facial expressions paired with angry sounds, $p = .433$. The pattern of the results is inconsistent with the prediction of emotion seeds view, without observing greater interference effects from high similarity sounds than from low similarity sounds when recognising surprised facial expressions.

4.2.2.2 Errors Analyses

The same error analyses as in experiment 1 were performed in the following subsections.

4.2.2.2a Analysis for Mis-categorisation of Angry Facial Expressions

Table 4.7 presents the percentage of mis-categorisations of angry facial expressions paired with disgusted sounds. A one-way repeated-measures ANOVA (emotion: disgust, fear, happiness, sadness, or surprise) was conducted on the mis-categorisation of angry facial expressions under the condition of high similarity. There was a main effect of emotion on the mis-categorisation of angry faces, $F(1.10, 25.25) = 37.51$, $p < .001$, $\eta^2_p = .620$. Pairwise comparisons (Fisher's LSD) revealed that the mis-categorisation as disgust (44.47%) was significantly higher than any other incongruent emotions, all *ps* < .001. The mis-categorisation of angry facial expressions as happy (0%) was significantly lower than that as fearful (2.61%), p = .005, as sad (4.70%), p = .005, and as surprised (0.63%), p = .031. The percentage of mis-categorisation of angry faces as surprised was significantly lower than that as fearful, p = .024, and lower than that as sad, p = .010. There was no significant difference between the mis-categorisation of angry faces as fearful and as sad, p = .203. The results are consistent with emotion seeds view which predicted that a facial expression paired with highly similar context is more likely to be mis-categorised as the emotion conveyed by highly similar context than as any other incongruent emotions.

Table 4.7. Mean Percentage of Mis-categorisation of Facial Expressions Paired with High Similarity Sounds (%)

Face-sound Composite	Miscategorisation					
	Anger	Disgust	Fear	Happiness	Sadness	Surprise
Anger-Disgust		44.47 (6.13)	2.61 (0.83)	0.00	4.70 (1.50)	0.63 (0.27)
Disgust-Anger	35.96 (5.26)		7.36 (1.69)	0.42 (0.24)	16.64 (3.97)	1.26 (0.60)
Fear-Surprise	3.23 (1.06)	4.17 (1.59)		0.31 (0.17)	3.44 (0.86)	38.62 (4.34)
Happiness-Surprise	1.35 (0.72)	0.42 (0.28)	1.98 (0.78)		0.94 (0.38)	26.35 (6.94)
Sadness-Fear	5.35 (1.13)	5.44 (1.45)	25.63 (5.61)	1.16 (0.42)		4.71 (1.30)
Surprise-Fear	1.98 (0.76)	0.52 (0.33)	38.24 (5.40)	3.13 (1.28)	0.83 (0.35)	

- SEM in Parentheses

The percentage of the mis-categorisation of angry faces as disgust across the conditions of perceptual similarity is listed in Table 4.8. A one-way repeated-measures ANOVA (perceptual similarity: isolated face, high similarity, low similarity) was conducted for the mis-categorisation of angry faces as disgust. A main effect of perceptual similarity was significant, $F(1.16, 26.63) = 28.21$, $p < .001$, $\eta^2_p = .551$. Pairwise comparisons (Fisher's LSD) showed that the percentage of mis-categorisation of angry faces paired with disgusted sounds (high similar context) as disgust (44.47%) was significantly more than that of angry faces paired with surprised sounds (low similar context) (9.06%) and isolated angry faces (9.58%), both *ps* < .001. And the percentage of mis-categorisation of angry faces paired with surprised sound as disgusted did not significantly differ from that of mis-categorisation of isolated angry faces as disgusted, *p* = .807. The results confirm that emotion seeds exert greater influence on recognition of angry faces when paired with disgusted sounds (that is, in the condition of high similarity context).

Table 4.8. Mean Percentage of Mis-categorisation of Facial Expressions as the Emotion by High Similarity Sounds across the Conditions of Perceptual Similarity (%)

Face-Miscategorisation	Perceptual Similarity		
	Isolated Face	High Similarity	Low Similarity
Anger-Disgust	9.58 (2.15)	44.47 (6.13)	9.06 (1.68)
Disgust-Anger	6.59 (1.72)	35.96 (5.26)	9.83 (2.05)
Fear-Surprise	34.17 (3.36)	38.62 (4.18)	20.65 (3.79)
Happiness-Surprise	0.21 (0.20)	26.35 (6.68)	0.73 (0.39)
Sadness-Fear	2.92 (1.17)	26.46 (5.40)	3.26 (1.01)
Surprise-Fear	16.25 (3.18)	38.24 (5.20)	24.77 (4.10)

- SEM in Parentheses

The separate paired-sample t-test for anger trials showed that the percentage of mis-categorisation as disgusted when angry facial expressions were paired with disgusted sounds (44.47%) was significantly greater than that of mis-categorisation as surprised when angry facial expressions were paired with surprised sounds (9.06%), $t(23) = 7.00$, $p < .001$, $d = 1.43$. The results are consistent with emotion seeds view, with that the percentage of target mis-categorisation was reduced as the level of perceptual similarity was decreased.

4.2.2.2b Analysis for Mis-categorisation of Disgusted Facial Expressions

Table 4.7 presents the percentage of mis-categorisations of disgusted facial expressions paired with angry sound. A one-way repeated-measures ANOVA (emotion: anger, fear, happiness, sadness, or surprise) was performed on the mis-categorisation of disgusted facial expressions under the condition of high similarity. There was a main effect of emotion on the mis-categorisation of disgusted faces, $F(1.68, 38.64) = 19.78$, $p < .001$, $\eta^2_p = .462$. Pairwise comparisons (Fisher's LSD) revealed that the mis-categorisation of disgusted facial expressions as anger (35.96%) was significantly higher than the mis-categorisations of disgusted faces as fearful (7.36%), happy (0.42%), and surprised (1.26%), all *ps* < .001, and higher than the mis-categorisation of disgusted faces as sad (16.64%), $p = .023$. The percentage of mis-categorisation of disgusted faces as sad was significantly higher than that of mis-categorisation of disgusted faces as happy and surprised, both *ps* < .001, and was marginally different from that of mis-categorisation of disgusted faces as fearful, p = .061. There was significant difference between the mis-categorisation of disgusted faces as fearful and as happy, p = .001, and significant difference between the mis-categorisation as fearful and as surprised, p = .004. The percentage of mis-categorisation of disgusted faces as happy was not significant different from that as surprised, p = .176. The results are consistent with emotion seeds view which predicted that a facial expression paired with highly similar context is more likely to be mis-categorised as the emotion conveyed by highly similar context than as any other incongruent emotions.

The percentage of mis-categorisation of disgust faces as anger across the conditions of perceptual similarity is listed in Table 4.8. A one-way repeated-measures ANOVA (perceptual similarity: isolated face, high similarity, low similarity) was conducted for the mis-categorisation of disgusted faces as anger. A main effect of perceptual similarity was significant, $F(1.28, 29.53) = 30.25$, $p < .001$, $\eta^2_p = .568$. Pairwise comparisons (Fisher's LSD) showed that the percentage of mis-categorisation of disgusted faces paired with angry sounds as anger (35.96%) was significantly more than that of mis-categorisation of disgusted faces paired with fearful sound as anger (9.83%) and that of mis-categorisation of isolated disgusted faces as anger (6.59%), both $ps < .001$. And the percentage of mis-categorisation of angry faces under the condition of low similarity as disgusted was not significantly different from that of mis-categorisation of isolated angry faces as disgusted, p = .135. The results confirm that emotion seeds exert greater influence on recognition of disgusted faces when paired with angry bodies (that is, in the condition of high similarity context).

The paired-sample t-test found that the percentage of mis-categorisation of disgusted faces with angry bodies as angry (35.96%) did not significantly differ from that of mis-categorisation of disgusted facial expressions with fearful bodies as fearful (29.76%), $t(23) = 1.29$, $p = 0.221$. The results are inconsistent with emotion seeds view which predicted greater percentage of target mis-categorisation elicited by high similarity sounds than by low similarity ones.

4.2.2.2c Analysis for Mis-categorisation of Fearful Facial Expressions

Table 4.7 presents the percentage of mis-categorisations of fearful facial expressions paired with surprised sounds. A one-way repeated-measures ANOVA (emotion: anger, disgust, happiness, sadness, or surprise) was conducted on the mis-categorisation of fearful facial expressions under the condition of high similarity (surprised sounds). There was a main effect of emotion on the mis-categorisation of fearful faces, $F(1.35, 31.12) = 49.99$, $p < .001$, $\eta^2_p = .685$. Pairwise comparisons (Fisher's LSD) showed that the mis-categorisation of fearful faces as surprised (38.62%) was significantly higher than any other incongruent emotions, all *ps* < .001. The mis-categorisation of fearful facial expressions as happy (0.31%) was significantly lower than as angry (3.23%), $p = .015$, as disgusted (4.17%), $p = .021$, and as sad (3.44%), $p = .002$. There were no significant differences between each pair of the mis-categorisation of fearful faces as angry, disgusted, and sad, all *ps* > .10. The results are consistent with emotion seeds view which predicted that a facial expression paired with highly similar context is more likely to be mis-categorised as the emotion conveyed by highly similar context than as any other incongruent emotions.

The percentage of mis-categorisation of fearful faces as surprise across the conditions of perceptual similarity is listed in Table 4.8. A one-way repeated-measures ANOVA (perceptual similarity: isolated face, high similarity, low similarity) was conducted for the mis-categorisation of fearful faces as surprise. The main effect of perceptual similarity was significant, $F(2, 46) = 9.02$, $p < .001$, $\eta^2_p = .282$. Pairwise comparisons (Fisher's LSD) showed that the percentage of mis-categorisation of fearful faces paired with surprised sounds (high similarity context) as surprise (38.62%) was significantly more than that of mis-categorisation of fearful faces paired with disgusted sound (low similarity context) as surprise (20.65%), $p = .002$, but was not significantly more than that of mis-categorisation of isolated fearful faces as surprise (34.17%), $p = .263$. The percentage of mis-categorisation of isolated fearful faces as surprise was significantly greater than that of mis-categorisation of fearful faces paired with disgusted sounds as surprise, $p = .003$. The results showed that there was no significant difference between the accuracy of fearful facial expression paired with high similarity context and that of isolated fearful faces. The findings are not consistent with emotion seeds view which predicts greater influence on recognition of fearful faces should occur when paired with surprised sound (that is, in the condition of high similarity context).

The paired-sample t-test for fear trials showed the percentage of mis-categorisation of fearful faces paired with surprised sound as surprised (38.62%) was not significantly different from that of mis-categorisation of fearful faces paired with disgusted sound as disgusted (47.08%), $t(23) = 1.15$, $p = 0.261$. The results are not consistent with emotion seeds view which predicted greater percentage of target mis-categorisation triggered by high similarity context than by low similar context.

4.2.2.2d Analysis for Mis-categorisation of Happy Facial Expressions

Table 4.7 presents the percentage of mis-categorisations of happy facial expressions paired with surprised sounds. The one-way repeated-measures ANOVA (emotion: anger, disgust, fear, sadness, or surprise) was conducted on the mis-categorisation of happy facial expressions under the condition of high similarity. There was a main effect of emotion on the mis-categorisation of happy faces, $F(1.03, 23.75) = 12.95$, $p = .001$, $\eta^2_p = .360$. Pairwise comparisons (Fisher's LSD) showed that the mis-categorisation of happy faces paired with surprised sounds as surprised (26.35%) was significantly higher than the mis-categorisation of happy faces as angry (1.35%), and as sad (0.94%), both *ps* = .002 and higher than that of happy faces as disgusted (0.42%), and as fearful (1.98%), both *ps* = .001. The mis-categorisation of happy faces as anger was marginally larger than that of happy faces as disgust, p = .083. The mis-categorisation of happy faces as sad did not significantly differ from that of happy faces as anger, disgust, and fear, all ps > .10. And the mis-categorisation of happy faces as anger was not significant different from that of happy faces as fear, p > .10. The results are consistent with emotion seeds view which predicted that a facial expression paired with highly similar context is more likely to be mis-categorised as the emotion conveyed by highly similar context than as any other incongruent emotions.

The percentage of mis-categorisation of happy faces as surprise across the conditions of perceptual similarity is listed in Table 4.8. A one-way repeated-measures ANOVA (perceptual similarity: isolated face, high similarity, low similarity) was conducted for the mis-categorisation of happy faces as surprise. A main effect of perceptual similarity was significant, $F(1.00, 23.09) = 13.67$, $p = .001$, $\eta^2_p = .373$. Pairwise comparisons (Fisher's LSD) showed that the percentage of mis-categorisation of happy faces paired with surprised sound as surprise (26.35%) was significantly more than that of mis-categorisation of happy faces paired with sad sound as surprise (0.73%) and that of mis-categorisation of isolated happy faces as surprise (0.21%), both ps = .001. The percentage of mis-categorisation of isolated happy faces as surprise was not significantly different from that of mis-categorisation of happy faces paired with sad sound as surprise, p = .170. The results are consistent with emotion seeds view, with that emotion seeds exert greater influence on recognition of happy faces when paired with surprised sound (that is, in the condition of high similarity context).

The paired-sample t-test for happy trials showed the percentage of mis-categorisation of happy faces paired with surprised sounds as surprised (26.35%) was significantly greater than that of mis-categorisation of happy faces paired with sad sounds as sad (15.00%), $t(23) = 2.09$, $p = .048$, $d = 0.427$. The results are consistent with emotion seeds view which predicted greater percentage of target mis-categorisation caused by high similarity context than by low similarity context.

4.2.2.2e Analysis for Mis-categorisation of Sad Facial Expressions

Table 4.7 presents the percentage of mis-categorisations of sad facial expressions paired with fearful sounds. A one-way repeated-measures ANOVA (emotion: anger, disgust, happiness, sadness, or surprise) repeated-measures ANOVA was conducted on the mis-categorisation of sad facial expressions under the condition of high similarity. There was a main effect of emotion on the mis-categorisation of sad faces, $F(1.23, 28.29) = 13.11$, $p = .001$, $\eta^2_p = .355$. Pairwise comparisons (Fisher's LSD) revealed that the mis-categorisation of sad faces as fearful (25.63%) was significantly more than that of sad faces as surprised (4.71%), $p < .001$, that of sad faces as disgusted (5.44%), $p = .003$, that of sad faces as angry (5.35%), $p = .002$, and that of sad faces as happy (1.16%), $p < .001$. The percentage of mis-categorisation of sad facial expressions as happy was significantly lower than that of mis-categorisation of sad faces as angry, $p = .002$, that of mis-categorisation of sad faces as disgusted, $p = .006$, and that of mis-categorisation of sad faces as surprised, $p = .024$. There were no significant differences between each pair of the mis-categorisation of sad faces as angry, as disgusted, and as surprised, all *ps* > .10. The results are consistent with emotion seeds view which predicted that a facial expression paired with highly similar context is more likely to be mis-categorised as the emotion conveyed by highly similar context than as any other incongruent emotions.

The percentage of mis-categorisation of sad faces as fear across the conditions of perceptual similarity is listed in Table 4.8. A one-way repeated-measures ANOVA (perceptual similarity: isolated face, high similarity, low similarity) was conducted for the mis-categorisation of sad faces as fear. A main effect of perceptual similarity was significant, $F(1.11, 25.63) = 15.84$, $p < .001$, $\eta^2_p = .408$. Pairwise comparisons (Fisher's LSD) showed that the percentage of mis-categorisation of sad faces paired with fearful sound as fear (25.63%) was significantly more than that of sad faces paired with happy sound as fear (3.26%) and that of isolated sad faces as fear (2.92%), both *ps* < .001. The percentage of mis-categorisation of sad faces as fear did not significantly differ between the condition of low similarity and isolated sad faces, $p = .826$. The results confirm that emotion seeds exert greater influence on recognition of sad faces when paired with fearful sound (that is, in the condition of high similarity context).

The paired-sample t-test for sadness trials showed that the mis-categorisation of sad faces paired with fearful sound as fearful (25.63%) was significantly greater than the mis-categorisation of sad faces paired with happy sound as happy (18.66%), $t(23) =$

1.26, p = .219. The results are inconsistent with emotion seeds view which predicted greater percentage of target mis-categorisation caused by high similarity context than by low similarity context.

4.2.2.2f Analysis for Mis-categorisation of Surprised Facial Expressions

Table 4.7 presents the percentage of mis-categorisations of surprised facial expressions paired with fearful sounds. A one-way repeated-measures ANOVA (emotion: anger, disgust, fear, happiness, or sadness) was conducted on the mis-categorisation of surprised facial expressions under the condition of high similarity. There was a main effect of emotion on the mis-categorisation of surprised faces, $F(1.16, 26.58) = 40.37$, $p < .001$, $\eta^2_p = .637$. Pairwise comparisons (Fisher's LSD) revealed that the mis-categorisation of surprised faces as fearful (38.24%) was significantly higher than any other incongruent emotions, all *ps* < .001. There were no other significant differences between each pair of the mis-categorisation as four other emotions (angry, 1.98%, disgusted, 0.52%, happy, 3.13%, and sad, 0.83%), except marginal difference between the mis-categorisation as disgusted and as happy, p = .074. The results are consistent with emotion seeds view which predicted that a facial expression paired with highly similar context is more likely to be mis-categorised as the emotion conveyed by highly similar context than as any other incongruent emotions.

The percentage of mis-categorisation of surprised faces as fear across the conditions of perceptual similarity is listed in Table 4.8. A one-way repeated-measures ANOVA (perceptual similarity: isolated face, high similarity, low similarity) was conducted for the mis-categorisation of surprised faces as fear. A main effect of perceptual similarity was significant, $F(1.48, 34.11) = 8.59$, $p = .002$, $\eta^2_p = .272$. Pairwise comparisons (Fisher's LSD) showed that the percentage of mis-categorisation of surprised faces paired with fearful sounds (high similarity context) as fear (38.24%) was significantly more than that of mis-categorisation of surprised faces paired with angry sounds (low similar context) as fear (24.58%), $p = .019$, and that of mis-categorisation of isolated sad faces as fear (16.25%), $p = .003$. The percentage of mis-categorisation of isolated surprised faces as fear was significantly lower than that of mis-categorisation of surprised faces paired with angry sound as fear, $p = .033$. The results confirm that emotion seeds exert greater influence on recognition of surprised faces when paired with fearful sounds (that is, in the condition of high similarity context).

The paired-sample t-test for surprise trials revealed that the percentage of mis-categorisation of surprised faces paired with fearful sounds as fearful (38.24%) was significantly greater than that of mis-categorisation of surprised faces paired with angry sounds as angry (18.10%), $t(23) = 3.90$, $p = .001$, $d = 0.79$. The results are consistent with emotion seeds view, with greater percentage of target mis-categorisation caused by high similarity context than by low similarity context.

4.2.3 Discussion

The accuracy of happy facial expression recognition was the best, and disgusted and fearful facial expressions were recognised the worst, which replicated the results in Experiment 1.

The results showed that the better performance of congruent face-sound stimuli compared to isolated faces was found for angry, disgust, and fearful faces. As found in Experiment 1, the recognition of happy and sad facial expressions was not improved by congruent happy and sad sounds. Besides, facilitation effects from congruent surprised sound were not observed for surprised facial expressions.

The main goal of Experiment 2 was to test whether the hypothesis of emotion seeds view is applicable to context effects from vocal expressions. Interference effects under both incongruent conditions (that is, the conditions of high similarity and low similarity) were found for the recognition of angry and sad facial expressions, with lower percentage of accurate recognition under high similarity vocal expressions than under low similarity ones. The findings are consistent with the prediction of emotion seeds view. Surprisingly, the recognition of happy faces paired with surprised sounds and happy faces paired with sad sounds was worse than that of isolated happy faces and that of happy faces paired with happy sounds, which was not observed when happy faces were paired with bodily expressions in Experiment 1. The results that larger interference effects from surprised sounds than those from sad sounds on recognition of happy faces are consistent with the prediction of emotion seeds view. The percentage of accurate recognition of surprised faces was lower under high similarity condition and low similarity condition, in contrast to surprised face in isolation and congruent surprised face-sound stimuli. However, there were no significant differences between surprised faces paired with high similar sounds and those paired with low similar sounds. As for disgusted and fearful faces, interference effects on the recognition of these two facial expressions were found under low similarity conditions, (that is, disgusted faces paired with fearful sound and fearful faces paired with disgusted sound). However, the percentages of accurate recognition of disgusted and fearful faces were not found under high similarity conditions, compared to isolated fearful faces. The pattern of context effects on the recognition of disgust and fear is not consistent with the hypothesis of emotion seeds views.

The analyses of accuracy showed that the pattern of vocal contextual effects on the recognition of angry, happy, and sad facial expressions support the hypothesis of emotion seeds view, with greater interference effects under the condition of high similarity sounds than under the condition of low similarity sounds. However, the pattern of context effects from vocal expressions on the recognition of three other facial expressions (surprised, fearful, and disgusted facial expressions) did not support the prediction of emotion seeds view in the light of accuracy analysis.

The analysis of mis-categorisation for each facial expression did not fully lend support to emotion seeds view, too. The same error analyses as Experiment 1 were conducted for all six basic facial expressions in Experiment 2. The error analysis carried on the emotional faces paired with high similarity sounds showed that all six facial expressions when paired with high similarity sounds were mis-categorised more as the emotions by the simultaneously presented high similarity sounds. The pattern of the results is consistent with emotion seeds view. When facial expressions are paired with high similarity sounds, facial expressions are more likely to be mis-categorised as the emotions conveyed by high similarity contexts.

A second error analysis was performed on mis-categorisation of facial expressions as the emotions by high similarity sounds across the conditions of high similarity sound, low similarity sound, and isolated faces. It was found that all six facial expressions, but fearful facial expressions, when paired with high similarity sounds, were mis-categorised more as the emotions by high similarity sounds, compared to facial expressions in isolation and facial expressions paired with low similarity sounds. For fearful facial expressions, they were mis-categorised as surprised more under the condition of isolated fearful faces and when they were paired with surprised sounds (high similarity context) than when they were paired with disgusted sounds (low similarity context). However, there was no significant difference of the mis-categorisation fearful faces as surprised between the condition of isolated fearful faces and the condition of fearful faces with surprised sound. The pattern of the results indicates that except fearful facial expressions, emotion seeds exert greater influence on the recognition of all five other facial expressions (anger, disgust, happiness, sadness, and surprise) when they were paired with the high similarity sounds.

The last error analysis showed that for angry, happy, and surprised facial expressions, the percentage of target mis-categorisation of the three facial expressions paired with high similarity sounds was greater than that of target mis-categorisation of the three facial expression paired with low similarity sounds. However, for disgusted, fearful, and sad facial expressions, there were no significant differences between the percentage of target mis-categorisation of the three facial expressions paired with high similarity sounds and that of target mis-categorisation of the three facial expressions paired with low similarity sounds. The findings are partially consistent with emotion seeds view which predicted that the percentage of mis-categorisation of facial expressions as the emotions by the concurrently presented context was reduced as the levels of perceptual similarity were decreased.

In summary, that congruent sounds are beneficial to the recognition of facial expressions was observed for angry, fearful, and disgusted facial expressions, but was not found for happy, sad, and surprised facial expressions. The prediction that larger interference effects from high similarity context than from low similarity context was supported by the results of accuracy analysis on angry, happy, and sad facial expressions. Moreover, the three error analyses showed that the pattern of mis-categorisation of emotional faces predicted by emotion seeds view was fully supported by the results of angry, happy, and surprised facial expressions. Taking accuracy and error analyses into account, only the pattern of context effects on the recognition of angry and happy facial expressions is consistent with the prediction of emotion seeds view.

4.3 Chapter General Discussion

Experiment 1 and 2 set out were to investigate whether the recognition of facial expressions was influenced by accompanied emotional context, and whether the pattern of context effects is consistent with the prediction of emotion seeds view. The current results from Experiment 1 and 2 confirm that the recognition of facial expressions is influenced by the concurrently presented bodily expressions and vocal expressions. The details will be discussed in the following section.

The accuracy and error analyses in experiment 1 revealed that the pattern of context effects from bodily expressions on the recognition of angry, sad, and surprised facial expressions is consistent with the prediction of emotion seeds view which proposed that larger context effects from high similarity bodies than from low similarity bodies. The pattern of context effects on the recognition of disgusted faces will be discussed later. The accuracy and error analyses in experiment 2 showed that the pattern of context effects from vocal expressions on the recognition of angry and happy facial expressions is consistent with the prediction of emotion seeds view. One possibility for the differences between experiment 1 and 2 is contextual stimuli used. The contextual stimuli in experiment 1 were bodily expressions which were static stimuli, whereas vocal expressions in Experiment 2 were dynamic stimuli. It was found that dynamic stimuli could provide more information than static ones (Grèzes, Pichon, & de Gelder, 2007), so that vocal expressions might be more influential than static bodily expressions. It might be the reason that the pattern of context effects from vocal expressions on recognition of facial expressions differed from those from bodily expressions on recognition of facial expressions.

In Experiment 1, unexpectedly, the accuracy of disgusted faces was significantly increased when disgusted faces paired with fearful bodies (low similarity context) compared to isolated disgusted faces. One possibility for the improvement for disgusted faces paired with fearful bodies is the similarity between emotional bodies. It was found that gestures of disgusted bodies are similar to those of fearful bodies (Wallbott, 1998), such as a whole body turned away, arms pressed close to the sides. The increase of the accuracy of disgusted faces paired with fearful bodies might because fearful bodily expressions look similar to disgusted ones. However, the similar result was not observed for fearful faces. The recognition of fearful faces was impaired when paired with disgusted bodies, rather than improved, compared to isolated fearful faces. This result to some extent is similar to that in the study carried out by Mondloch (2012) where she found that incongruent face-body stimuli (sad face-happy body and happy face-sad body) were recognised more accurately than congruent face-body stimuli (sad face-sad body and happy face-happy body). Mondloch (2012) did not provide any explanation, instead suggested more studies might be needed to find out whether the result in her study would be replicated.

Collectively, although there is not enough evidence to fully support the prediction of emotion seeds view, the current results indicate the emotional contexts, including bodily and vocal expressions, does play a role in the recognition of facial expressions. Congruent bodily and vocal expressions could be beneficial to recognition of facial expressions, while incongruent bodily and vocal expressions impair facial expression recognition to different degrees.

Chapter 5 Subtle Facial Expressions and Context Effects

In chapter 4 it has been demonstrated that accompanied bodily expressions and vocal expressions exert an influence on recognition of unambiguous facial expressions (namely, 100% intensity of facial expressions). And previous studies (Dolan et al., 2001; Lee et al., 2012; Müller et al., 2011) have found that besides unambiguous facial expressions, contextual information also has an impact on recognition of other levels of intensity of facial expressions (e.g., 40% intensity and 50% intensity of fearful facial expressions).

As reviewed in the introduction, context effects depend on the levels of ambiguity of emotional facial expressions (Trope, 1986). An ambiguous facial expression that Trope meant is a mixture of two emotional faces, such as, a mixture of happy-angry face, which I have termed Type A. It has been demonstrated that larger context effects occur when recognising more ambiguous facial expressions (van den Stock et al., 2007; de Gelder & Vroomen, 2000). For example, context effects on the recognition of ambiguous facial expressions (10%happiness-90%anger, or 40%happiness-60%anger) were increased compared to those on the recognition of unambiguous facial expressions (100%happiness, or 100%anger).

When it comes to the relationship between the magnitude of context effects and different levels of intensity of facial expressions, it is expected that contextual information would exert greater influence on recognising lower levels of intensity of facial expressions than higher levels of intensity of facial expressions. However, as there is a lack of studies on this topic, I used two experiments to explore whether the observed context effects in chapter 4 are influenced by different levels of intensity of facial expression. It is anticipated that greater context effects will be occurred on lower level of intensity of facial expressions.

As discussed in experiment 1, fearful and disgusted facial expressions can signal vital cues of danger in the environment, though they are the most difficult facial expressions to be recognised. Another common occurrence is that fearful and disgusted facial expressions themselves present avoidance reactions to danger and contamination in the environment, so as to alert other individuals to keeping away from danger. So in following chapters, I focused on the facial expressions of fear and disgust.

5.1 Experiment 3

This experiment explored whether the observed visual context effects are influenced by levels of intensity of facial expressions. It was expected that context effects would be greater on recognition of lower level of intensity of facial expressions.

5.1.1 Method

5.1.1.1 Participants

Thirty-two new participants between 18 to 32 years of age (15 females and 17 males) were recruited to take part in Experiment 3. Three participants were left-handed and the rest were right-handed. They all had given informed consent and were paid NZD 10 for participation. Half of the participants were arranged in the disgusted-face block, and half in the fearful-face block. All had normal or corrected-to-normal vision.

5.1.1.2 Apparatus and Stimuli

The apparatus used was the same as that described in Chapter 2.

The images of disgusted facial expressions and fearful facial expressions (2 females and 2 males) used were those in Experiment 3. The images of neutral faces were taken from KDEF set (Lundqvist et al., 1998) and these images were from the same identities displaying disgusted and fearful facial expressions used in this experiment. I used *Abrosoft FantaMorph 5* to generate images of 3 levels of intensity of facial expressions (including 40%, 60%, and 80%) from morphing neutral faces to disgusted (or, fearful) facial expressions of the same identity. Unambiguous disgusted and fearful faces (that is, the level of 100% intensity) were also used in the experiment.

The images of bodily expressions of disgust, anger, fear, and surprise, which were the same as those in Experiment 1, were used in this experiment. I used Adobe Photoshop CS5 (trial version) to create congruent-disgusted and incongruent-disgusted pairs by pasting the images of different levels of intensity of disgusted facial expressions on the images of disgusted bodies and angry bodies, and to create congruent-fearful and incongruent-fearful pairs by pasting the images of different levels of intensity of fearful facial expressions on the images of fearful bodies and surprised bodies. The reason for the choice of angry bodies and surprised bodies to create incongruent-disgusted face-body stimuli and incongruent-fearful face-body stimuli was that the recognition of facial expressions of disgust and fear was recognised the worst when disgusted facial expressions were paired with angry bodies and fearful facial expressions were paired with surprised bodies, in experiment 1. Isolated faces served as baseline stimuli and were presented in a separate block.

5.1.1.3 Design and Procedure

This experiment used a 2 (facial expression: disgust or fear) x 3 (consistency condition: isolated faces, congruent condition, or incongruent condition) x 4 (level of intensity: 40%, 60%, 80%, or 100%) mixed factorial design. The factor of facial expression was the between-subjects factor, and the factors of level of intensity and consistency condition were the within-subjects factors.

The procedure was identical to that described in chapter 2, except for the choices of facial expressions. In this experiment, a response screen with two options (with disgust and anger in disgust block, and fear and surprise in fear block), instead of six options, was shown to participants until judgment was made or disappeared after 8 s without any response. The order of the two options was balanced across participants. Participants were asked to use the mouse to click on one of the two options to indicate expressions on the faces when presented in isolation block, and when in face-body compounds block they needed to judge expressions on the faces while ignoring bodily expressions. The experiment in each block consisted of 600 trials, including 10 repetitions of 20 congruent trials, 20 incongruent trials, and 20 trials of isolated faces. The order of isolated face block and face-body compounds block was balanced across participants. All participants had a rest every 200 trials. Practice trials were given to participants to make them familiar with the experiment. The trials used for practice were not used in the main test.

5.1.2 Results

The goal of this experiment was to examine the relationship between visual context effects and the levels of intensity of facial expressions.

The results of experiment 3 are plotted in Figure 5.1. A 3 x 4 x 2 mixed ANOVA with consistency condition (face-only condition, congruent condition, or incongruent condition) and level of intensity (40%, 60%, 80%, or 100%) as within-subjects factors and with facial expression (disgust or fear) as a between-subjects factor was performed. A main effect of consistency condition was found, $F(1.39, 41.67) = 45.75$, $p < .001$, $\eta^2_p = .604$. Congruent face-body stimuli (77.85%) were recognised better than isolated faces (56.23%) and incongruent ones (37.07%), both *ps* < .001. The accuracy of incongruent face-body stimuli was significantly lower than that of isolated faces, $p < .001$. A significant main effect of level of intensity was observed, $F(1.17, 35.17) = 15.74$, $p < .001$, $\eta^2_p = .344$. The accuracy for the level of intensity of 40% (46.12%) was significantly lower than that for level of 60% intensity (55.63%), $p = .001$, and was significantly less than that for levels of 80% intensity (62.01%) and 100% intensity (64.45%), both *ps* < .001. The accuracy for the level of intensity of 60% was significantly lower than that for the level of 80% intensity, $p = .001$, and lower than that for the level of 100%, $p < .001$. The accuracy for the level of 80% intensity was significantly lower than that for the level of 100% intensity, $p = .024$. A main effect of facial expression was significant, $F(1, 30) = 4.31$, $p = .046$, $\eta^2_p = .126$, with the accuracy of disgusted facial expressions (61.91%) higher than the accuracy of fearful facial expressions (52.19%).

The two-way interaction between facial expression and consistency condition was significant, $F(1.39, 41.67) = 6.34$, $p = .008$, $\eta^2_p = .177$. For disgusted facial expressions, the congruent face-body stimuli (88.44%) were significantly recognised better than the incongruent face-body stimuli (33.24%) and isolated disgusted faces (39.37%), both *ps*

< .001. The accuracy of incongruent face-body stimuli was significantly lower than that of isolated faces, $p < .001$. For fearful facial expressions, the accuracy of congruent face-body stimuli (67.27%) was significantly higher than that of incongruent face-body stimuli (40.90%) and higher than that of isolated fearful faces (48.40%), both *ps* = .002. The accuracy of incongruent face-body stimuli was marginally lower than that of isolated fearful faces, $p = .089$.

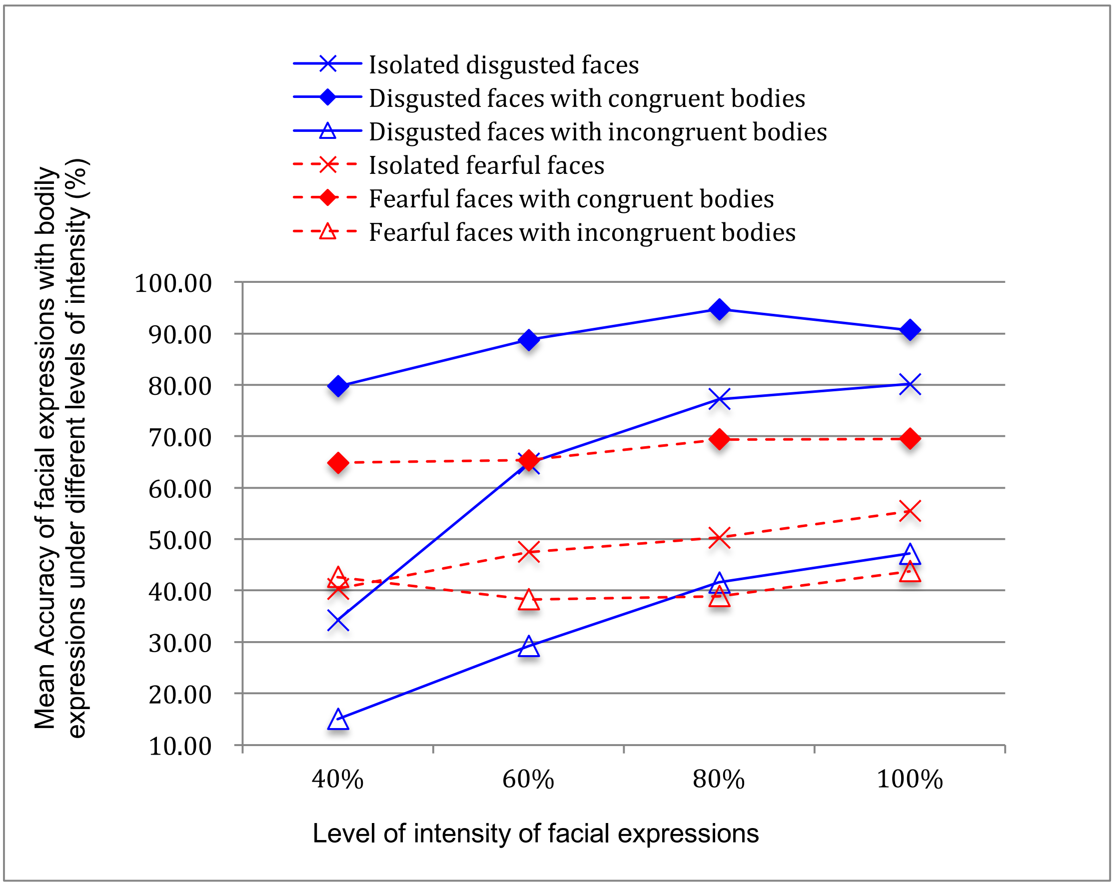

Figure 5.1. The Results from Experiment 3.

The two-way interaction between facial expression and level of intensity, $F(1.17, 35.17) = 7.35$, $p = .008$, $\eta^2_p = .197$. For disgusted facial expressions, the accuracy of the level of 40% intensity (42.97%) was significantly lower than that of levels of 60% intensity (60.89%), 80% intensity (71.15%) and 100% intensity (72.66%), all *ps* $< .001$. The accuracy of level of 60% intensity was significantly lower than that of level of 80% intensity, $p < .001$, and lower than level of 100% intensity, $p = .001$. And there was no significant difference of the accuracy between the level of 80% intensity and the level of 100% intensity, $p = .309$. For fearful facial expressions, the accuracy was reduced as the level of intensity was decreased: from level of 100% intensity (56.25%) to level of 80% intensity (52.87%) to level of 60% intensity (50.37%) to level of 40% intensity (49.27%). The accuracy of the level of 60% intensity was marginally lower than that of the level of 100% intensity, $p = .063$. Fearful facial expressions at the level of 80% intensity were recognised significantly worse than those at the level of 100% intensity, $p = .028$. There were no other significant differences between each pairs of other levels of intensity, all *ps* $> .10$.

The two-way interaction between consistency condition and level of intensity was significant, $F(3.13, 93.84) = 10.93$, $p < .001$, $\eta^2_p = .267$. Under the level of 40% intensity, the accuracy of congruent face-body stimuli (72.27%) was significantly greater than that of incongruent face-body stimuli (28.83%) and isolated faces (37.27%), both *ps* $< .001$. The accuracy of incongruent face-body stimuli was significantly lower than that of isolated faces, $p = .042$. Under the level of 60% intensity, the accuracy of congruent face-body stimuli (77.03%) was significantly greater than that of incongruent face-body stimuli (33.75%) and that of isolated faces (56.09%), both *ps* $< .001$. The accuracy of incongruent face-body stimuli was significantly lower than that of isolated faces, $p < .001$. Under the level of 80% intensity, the accuracy of congruent face-body stimuli (82.03%) was significantly higher than that of incongruent face-body stimuli (40.23%) and that of isolated faces (63.75%), both *ps* < .001. The incongruent face-body stimuli were recognised worse than the isolated faces, $p < .001$. Under the level of 100% intensity, the accuracy of congruent face-body stimuli (80.08%) was significantly greater than that of incongruent face-body stimuli (45.47%), $p < .001$, and greater than that of the isolated faces (67.81%), $p = .004$. And, the accuracy of incongruent face-body stimuli was significantly lower than that of the isolated faces, $p < .001$.

The three-way interaction was significant, $F(3.13, 93.84) = 4.24$, $p = .007$, $\eta^2_p = .124$. And, the three-way interaction was plotted in Figure 5.1. For disgusted faces under the level of 40% intensity, the accuracy of congruent face-body stimuli (79.69%) was significantly greater than that of incongruent face-body stimuli (15%) and greater than that of isolated disgusted faces (34.22%), both *ps* < .001. The accuracy of incongruent face-body pairs was significantly lower than that of isolated faces, $p = .002$. For disgusted faces under the level of 60% intensity, the accuracy of congruent face-body stimuli (88.75%) was significantly greater than that of incongruent face-body pairs (29.22%) and that of isolated faces (64.69%), both *ps* < .001. The accuracy of incongruent face-body stimuli was significantly higher than that of isolated faces, $p < 0.001$. For disgusted faces under the level of 80% intensity, the accuracy of congruent face-body stimuli (94.69%) was significantly greater than that of incongruent face-body stimuli (41.56%), $p < .001$, and greater than that of isolated disgusted faces (77.19%), $p = .001$. The accuracy of incongruent face-body stimuli was significantly lower than that of isolated faces, $p < .001$. For disgusted faces under the level of 100% intensity, the recognition of congruent face-body stimuli (90.63%) and isolated disgusted faces (80.16%) was significantly better than that of incongruent face-body stimuli (47.19%), both *ps* < .001. There were marginally significant differences between the congruent face-body stimuli and isolated faces, $p = .066$.

For fearful faces under the level of 40% intensity, the accuracy of congruent face-body stimuli (64.84%) was significantly greater than that of incongruent face-body stimuli (42.66%), $p = .012$, and greater than that of isolated fearful faces (40.31%), $p = .008$. The accuracy of incongruent face-body stimuli did not significantly differ from that of isolated faces, $p = .679$. For fearful faces under the level of 60% intensity, the accuracy of congruent face-body stimuli (65.31%) was significantly different from that of incongruent face-body stimuli (38.28%), $p = .002$, and significantly differed from that of isolated fearful faces (47.50%), $p = .005$. The accuracy of incongruent face-body stimuli was marginally lower than that of isolated fearful faces, $p =0.092$. For fearful faces under the level of 80% intensity, the accuracy of congruent face-body stimuli (69.38%) was significantly greater than that of incongruent face-body stimuli (38.91%), $p < .001$, and greater than that of isolated fearful faces (50.31%), $p = .001$. The accuracy of incongruent face-body stimuli was significantly lower than that of isolated fearful faces, $p = .024$. For fearful faces under the level of 100% intensity, the accuracy of congruent face-body stimuli (69.53%) was significantly higher than that of incongruent face-body stimuli (43.75%), $p = .004$, and higher than that of isolated fearful faces (55.47%), $p = .016$. The accuracy of incongruent face-body stimuli was significantly lower than that of isolated fearful faces, $p = .030$.

The three-way interaction can be decomposed in another way. At the level of 40% intensity, there were no significant differences between the accuracy of isolated disgusted facial expressions and that of isolated fearful ones, $p = .557$. Isolated disgusted facial expressions were recognised significantly better than isolated fearful ones at three other levels of intensity (with $p = .014$ at the level of 60% intensity, $p < .001$ at the level of 80% intensity, and $p < .001$ at the level of 100% intensity). At the level of 40%, there was no significant difference between the accuracy of congruent disgusted face-body stimuli and that of congruent fearful face-body stimuli, $p = .131$. Like faces presented in isolation, congruent disgusted face-body stimuli were recognised significantly better than congruent fearful face-body stimuli at three other levels of intensity (with $p = .004$ at the level of 60% intensity, $p < .001$ at the level of 80% intensity, and $p = .002$ at the level of 100% intensity). In contrast, at the level of 40% intensity, incongruent disgusted face-body stimuli were recognised significantly worse than incongruent fearful face-body stimuli, $p = .007$. There was no significant difference between incongruent disgusted face-body stimuli and incongruent fearful face-body stimuli at three other levels of intensity, all *ps* > .10.

5.1.2.1 The Analysis for the Magnitude of Visual Consistency Effects on Different Levels of Intensity of Facial Expressions

The main interest of this experiment was to directly test the relationship between visual consistency effects and the level of intensity of facial expressions. The consistency effects refer to the differences of the accuracy between congruent face-body stimuli and incongruent face-body stimuli. In order to compare the results of this current experiment with those of previous studies, I have performed an analysis for the magnitude of consistency effects on four levels of intensity of facial expressions. The consistency effects, here I mentioned, are the differences of accuracy between congruent face-body stimuli and incongruent face-body stimuli.

The visual consistency effects for each level of intensity list in Table 5.1. A 2 (facial expression: disgust or fear) x 4 (level of intensity: 40%, 60%, 80%, or 100%) mixed ANOVA was carried out on consistency effects, with facial expression as a between-subjects factor and level of intensity as a within-subjects factor. A main effect of facial expression was significant, $F(1, 30) = 7.01$, $p = .013$, $\eta^2_p = .189$, with larger consistency effects for disgusted facial expressions (55.20) than those for fearful facial expressions (26.37). A main effect of level of intensity was significant, $F(2.24, 67.10) = 6.20$, $p = .002$, $\eta^2_p = .171$. The consistency effects for level of 100% intensity (34.61) was significantly less than that for level of 40% intensity (43.44), $p = .006$, less than that for level of 60% intensity (43.28), p = .002, and less than that for level of 80% intensity (41.80), p < .001.

Table 5.1. Visual Consistency Effects across Different Levels of Intensity of Facial Expressions

	Intensity of facial expressions			
Facial Expression	40%	60%	80%	100%
Disgust	64.69	59.53	59.13	34.61
Fear	22.19	27.03	30.47	25.78

The interaction between facial expression and level of intensity was significant, $F(2.24, 67.10) = 10.74$, $p < .001$, $\eta^2_p = .264$. For disgusted facial expressions, consistency effects under level of 100% intensity (34.61) were significantly less than those under level of 40% intensity (64.69), less than those under level of 60% intensity (59.53), and less than those under level of 80% intensity (59.13), all *ps* < .001. Consistency effects under level of 80% intensity was significantly less than those under level of 40% intensity, p = .003, and less than those under level of 60% intensity, p = .039. There was no significant difference between consistency effects under level of 40% intensity and those under level of 60% intensity, p > .10. For fearful facial expressions, consistency effects under level of 100% intensity (25.78) was marginally less than those under level of 80% intensity (30.47), p = .05, but did not significantly differ from those under level of 40% intensity (22.19) and those under level of 60% intensity

(27.03), both *ps* > .10. Consistency effects under level of 80% intensity was significantly greater than those under level of 40% intensity, p = .029, but did not significantly differ from those under level of 60% intensity, p > .10. There was no significant difference between consistency effects under level of 40% intensity and those under level of 60% intensity, p > .10.

The following analyses were carried out separately on facilitation effects and interference effects on the recognition of facial expressions across different levels of intensity in the following two subsections. Facilitation effects were obtained from subtracting the accuracy of face-only condition from the accuracy of congruent face-body condition. Interference effects were obtained from subtracting the accuracy of face-only condition from the accuracy of incongruent face-body condition.

5.1.2.2 The Analysis for the Magnitude of Visual Facilitation Effects on Different Levels of Intensity

The data for facilitation effects presents in Table 5.2. A separate 4 x 2 mixed ANOVA with level of intensity (40%, 60%, 80%, or 100%) as a within-subjects factor and with facial expression (disgust or fear) as a between-subjects factor was conducted. A main effect of level of intensity for facilitation effects was found, $F(1.80, 53.90) = 15.64$, $p < .001$, $\eta^2_p = .343$. The greatest facilitation effects were found for the lowest intensity (40%) of facial expressions (35.00), which was significantly larger than 60% level of intensity (20.94), p = .007, 80% level of intensity (18.28), p = .004, and 100% level of intensity (12.27), $p < .001$. The medium facilitation effects were observed under 60% level of intensity, and 80% level of intensity, without any significant difference between each other, $p > .10$. The facilitation effects under 100% level of intensity were significantly lower than those under 60% level of intensity and lower than those under 80% intensity, p = .023 and p = .014, respectively. There was no significant main effect of facial expression, $F(1, 30) = 0.48$, p = . 494.

A significant interaction between level of intensity and facial expression was found, $F(1.80, 53.90) = 5.22$, p = .011, $\eta^2_p = .148$. For disgusted facial expressions, facilitation effects were decreased from lower level of intensity to higher level of intensity: 45.47 at level of 40% intensity, 24.06 at level of 60% intensity, 17.50 at level of 80% intensity, and 10.47 at level of 100% intensity. Pairwise comparisons (Fisher's LSD) showed that the magnitude of facilitation effects under level of 40% intensity was significantly larger than that of facilitation effects under level of 60% intensity, p = .001, and larger than that of facilitation effects under levels of 80% intensity and 100% intensity, both *ps* $< .001$. The facilitation effects for level of 60% intensity were significantly greater than those for level of 100% intensity, p = .002 and were marginally greater than those for level of 80% intensity, p = .062. The magnitude of facilitation effects for level of 80% intensity was significantly greater than that of facilitation effects for level of 100% intensity, p = .010. For fearful facial expressions, facilitation effects under level of 40% intensity was 24.53, under level of 60% intensity was 17.81, under level of 80%

intensity was 19.06, and under level of 100% intensity was 14.06. It was only found a marginally difference of facilitation effects between level of 80% intensity and level of 100% intensity, *p* = .059. There was no other significant difference of facilitation effects of fearful facial expressions between any two levels of intensity, all *ps* > .10.

Table 5.2. Visual Facilitation Effects on Different Levels of Intensity of Facial Expressions

	Level of Intensity			
Facial Expression	40%	60%	80%	100%
Disgust	45.47	24.06	17.50	10.47
Fear	24.53	17.81	19.06	14.06

5.1.2.3 The Analysis for the Magnitude of Visual Interference Effects on Different Levels of Intensity

The data for interference effects presents in Table 5.3. A 4 x 2 mixed ANOVA with level of intensity (40%, 60%, 80%, and 100%) as a within-subjects factor and with facial expression (disgust and fear) as a between-subjects factor was conducted. A main effect of level of intensity for interference effects was found, $F(2.01, 60.42) = 8.51$, $p = .001$, $\eta^2_p = .221$. Pairwise comparisons (Fisher's LSD) showed that interference effects were smallest under level of 40% intensity (-8.44), which was significantly less than that under level of 60% intensity (-22.35) and under level of 80% intensity (-23.52), both *ps* =0.001, and significantly less than that under level of 100% intensity (-22.35), $p = .004$. There were no other significant differences between each pair of three other levels of intensity, all *ps* > .10. A main effect of facial expression for interference effects was found, $F(1,30) = 14.89$, $p = .001$, $\eta^2_p = .332$, with larger interference effects for disgusted face (-30.82) than that for fearful face (-7.5). There was no significant interaction between level of intensity and facial expression, $F(2.01, 60.42) = 0.23$, $p = .872$.

Table 5.3. Visual Interference Effects on Different Levels of Intensity of Facial Expressions

	Level of Intensity			
Facial Expression	40%	60%	80%	100%
Disgust	-19.22	-35.47	-35.63	-32.97
Fear	2.34	-9.22	-11.41	-11.72

5.1.3 Discussion

The experiment set out to examine the relationship between consistency effects from bodily expressions and levels of intensity of facial expressions. The larger consistency effects would be expected if the level of intensity of a facial expression were low, and would reduce if level of intensity increased. The predictions were based on previous studies (Trope, 1986; van den Stock et al., 2007; de Gelder & Vroomen, 2000), which demonstrated consistency effects depend on the level of ambiguity of a type A ambiguous facial expression, with larger consistency effects for an ambiguous facial expression than an unambiguous facial expression. For example, the magnitude of consistency effects is larger for 20%happiness-80%surprise than for 100%suprise.

The results from the current study revealed that the pattern of visual consistency effects (differences between congruent pairs and incongruent pairs) on the recognition of disgusted and fearful facial expressions across different levels of intensity was not exactly the same. The pattern of visual consistency effects on disgusted facial expression recognition is in line with the hypothesis, with larger consistency effects for level of 40% intensity that is the lowest level of intensity used in the current study. For the pattern of visual consistency effects on the recognition of fearful facial expressions, the larger consistency effects were observed when fearful faces under level of 80% intensity than unambiguous fearful faces (level of 100% intensity).

Moreover, consistency effects were divided into two types of effects, that is, facilitation effects and interference effects. The pattern of magnitude of facilitation effects on facial expressions of disgust and fear extends the prediction that larger consistency effects under lower level of intensity. The results showed that facilitation effects on the recognition of disgusted facial expressions was largest under level of 40% intensity (the lowest intensity used in this experiment), and was lowest under level of 100% intensity. In regard to facilitation effects on the recognition of fearful facial expressions, a marginally larger facilitation effect was found on level of 80% intensity than on level of 100% intensity.

As for interference effects, the pattern of interference effects showed the opposite tendency to that of facilitation effects. It was found that the pattern of interference effects on the recognition of disgusted facial expressions was identical to that of interference effects on recognition of fearful ones. The magnitude of interference effects for fearful faces and disgusted ones depends on the level of intensity of facial expressions. Interference effects observed on facial expressions of disgust and fear was smallest at the lowest level of intensity (40%). And interference effects became larger when the level of intensity of facial expression was increased, but without significant differences between each pair of levels of 60% intensity, 80% intensity and 100% intensity.

5.2 Experiment 4

This experiment explored whether the observed auditory context effects are influenced by levels of intensity of facial expressions. It was expected, like the results in experiment 3, that auditory context effects would be greater in lower level of intensity of facial expressions.

5.2.1 Method

5.2.1.1 Participants

Twenty-three new participants between 18 to 31 years of age (12 females and 11 males) were recruited to take part in Experiment 4. Two participants were left-handed and 21 were right-handed. Twelve participants were arranged in the disgusted-face block, and 11 in the fearful-face block. They all had given informed consent and were paid NZD 10 for their participation. All had normal/corrected-to-normal vision and normal hearing ability.

5.2.1.2 Apparatus and Stimuli

The apparatus was the same as that described in chapter 2.

The images of disgusted and fearful face images in Experiment 3 were used in this experiment. The affective bursts of disgust and fear were the same as those in Experiment 2. Face images and affective bursts were presented simultaneously to create emotional-congruent or emotional-incongruent pairs. The reason for the choice of disgusted sound and fearful sound to create incongruent-disgusted stimuli and incongruent-fearful ones was that the recognition of disgusted faces paired with fearful sounds and fearful faces paired with disgusted sounds was the worst in experiment 2. Isolated faces served as baseline stimuli and were presented in separate block.

5.2.1.3 Design and Procedure

This experiment used a 2 (facial expression: disgust or fear) x 4 (level of intensity: 40%, 60%, 80%, or 100%) x 3 (consistency condition: isolated faces, congruent condition, or incongruent condition) mixed factorial design. The factor of facial expression was the between-subjects factor, and the factors of level of intensity and consistency condition were the within-subjects factors.

The procedure was almost the same to that described in Experiment 3, except participants were asked to judge the expressions on faces while ignoring the accompanied sounds when presented with face-sound stimuli. The response screen presented two options (disgust and fear) to participants in the disgusted face block and the fearful-face block.

5.2.2 Results

The goal of this experiment was to examine the relationship between the auditory context effects and the levels of intensity of facial expressions.

The results of experiment 4 are plotted in Figure 5.2. A 3 x 4 x 2 mixed ANOVA with consistency condition (face-only condition, congruent condition, or incongruent condition) and level of intensity (40%, 60%, 80%, or 100%) as within-subjects factors and with facial expression (disgust or fear) as a between-subjects factor was performed. The analysis showed a main effect of consistency condition, $F(1.18, 24.74) = 24.16$, $p < .001$, $\eta^2_p = .535$, with higher accuracy on congruent face-sound stimuli (76.34%) than incongruent ones (37.05%) and face-only condition (57.38%), both *ps* < .001, and significant differences between incongruent face-sound stimuli and face-only condition, $p = .001$. A main effect for level of intensity was observed, $F(1.23, 25.91) = 9.96$, $p = .002$, $\eta^2_p = .322$. The accuracy of the level of 40% intensity (43.76%) was significantly lower than that of level of intensity of 60% intensity (57.38%), $p = .002$, lower than that of level of 80% intensity (62.01%), $p = .005$, and lower than that of level of 100% intensity (64.53%), $p = .002$. The accuracy of level of 60% intensity was marginally less than that of level of 80% intensity, $p = .076$, and was significantly less than that of level of 100% intensity, $p = .030$. The accuracy of level of 80% intensity was marginally less than that of level of 100% intensity, $p = .085$. A main effect of facial expression was not significant, $F(1, 21) = 0.491$, $p = .491$, with accuracy rates of disgusted faces being 58.39% and accuracy rates of fearful faces being 55.46%.

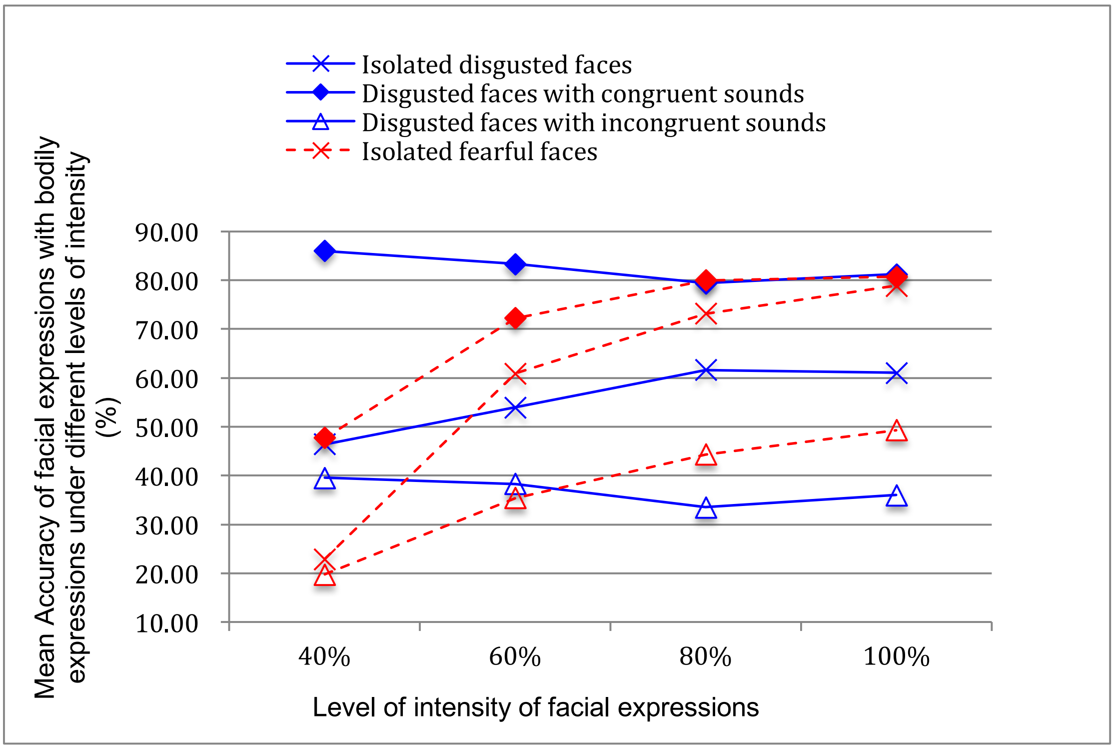

Figure 5.2. The Results from Experiment 4.

The two-way interaction between facial expression and level of intensity was significant, $F(1.23, 25.91) = 8.45$, $p = .005$, $\eta^2_p = .287$. For disgusted facial expressions, there were no significant differences between each pairs of the accuracy of level of 40% intensity (57.36%), level of 60% intensity (58.54%), level of 80% intensity (58.19%), and level of 100% intensity (59.44%), all *ps* > .10. For fearful facial expressions, the accuracy of the level of 40% intensity (30.15%) was significantly less than that of the level of 60% intensity (56.21%), less than that of the level of 80% (65.83%), and less than that of the level of 100% (69.62%), all *ps* < .001. The accuracy of the level of 60% intensity was significantly greater than that of the level of 80% intensity, $p = .014$, and significantly greater than that of the level of 100% intensity, $p = .007$. The accuracy of the level of 80% intensity was marginally less than that of the level of 100% intensity, $p = .073$.

The two-way interaction between consistency condition and level of intensity was significant, $F(1.74, 36.54) = 7.04$, $p = .004$, $\eta^2_p = .251$. Under the level of 40% intensity, the accuracy of congruent face-sound stimuli (66.88%) was significantly greater than that of incongruent face-sound stimuli (29.68%) and isolated faces (34.71%), both *ps* < .001. The accuracy of incongruent face-sound stimuli was not significantly different from that of isolated faces, $p = .17$. Under the level of 60% intensity, the accuracy of congruent face-sound stimuli (77.80%) was significantly greater than that of incongruent face-sound stimuli (36.89%) and that of isolated faces (57.43%), both *ps* < .001. The accuracy of incongruent face-sound stimuli was significantly lower than that of isolated faces, $p < .001$. Under the level of 80% intensity, the accuracy of congruent face-sound stimuli (79.69%) was significantly higher than that of incongruent face-sound stimuli (38.93%), $p < .001$, and was significantly higher than that of isolated faces (67.42%), $p = .007$. The incongruent face-sound stimuli were worse accurately recognised than the isolated faces, $p < .001$. Under the level of 100% intensity, the accuracy of congruent face-sound stimuli (80.97%) was significantly greater than that of incongruent face-sound stimuli (42.68%), $p < .001$, and was significantly greater than that of the isolated faces (69.95%), $p = .001$. And, the accuracy of incongruent face-sound stimuli was significantly lower than that of the isolated faces, $p = .002$.

The two-way interaction between facial expression and consistency condition was non-significant, $F(1.18, 24.74) = 1.07$, $p = .325$. The three-way interaction was not significant, $F(1.74, 36.54) = 0.53$, $p = .563$.

5.2.2.1 The Analysis for the Magnitude of Auditory Consistency Effects on Different Levels of Intensity of Facial Expressions

The main interest in experiment 4 was to directly test the relationship between the magnitude of auditory consistency effects and the level of intensity of facial expressions. Consistency effects for each level of intensity list in Table 5.4. A 2 (facial expression: disgust or fearful) x 4 (level of intensity: 40%, 60%, 80%, or 100%) mixed ANOVA was carried out on consistency effects (difference between congruent face-sound stimuli and incongruent face-sound stimuli), with facial expression as a between-subjects factor and level of intensity as a within-subjects factor. A main effect of facial expression was non-significant, $F(1, 21) = 0.69$, $p = .24$, with consistency effects for disgusted facial expressions being 45.63 and those for fearful facial expressions being 32.96. A main effect of level of intensity was non-significant, $F(3, 63) = 0.95$, $p = .42$. The interaction between facial expression and level of intensity was non-significant, $F(3, 63) = 1.45$, $p = .41$.

Table 5.4. Auditory Consistency Effects on Different Levels of Intensity of Facial Expressions.

Facial Expression	Level of Intensity			
	40%	60%	80%	100%
Disgust	46.46	45.00	45.83	45.63
Fear	27.96	36.82	35.68	31.36

The following analyses were carried out for facilitation effects and interference effects.

5.2.2.2 The Analysis for the Magnitude of Auditory Facilitation Effects on Different Levels of Intensity of Facial Expressions

The data for auditory facilitation effects presents in Table 5.5. A 4 x 2 mixed ANOVA with level of intensity (40%, 60%, 80%, and 100%) as a within-subjects factor and with facial expression (disgust and fear) as a between-subjects factor was conducted. A main effect of level of intensity for facilitation effects was found, $F(1.34, 28.15) = 7.21$, $p = .007$, $\eta^2_p = .256$. The largest facilitation effects were found for the lowest intensity (40%) of facial expressions (32.18%), which was significantly larger than 60% intensity (20.37%), $p = .009$, 80% intensity (12.26%), $p = .009$, and 100% intensity (11.01%), $p = .009$. The medium facilitation effects were observed under 60% intensity, with a marginal difference from 80% intensity, $p = .051$ and with a significant difference from 100% intensity, $p = .038$. Facilitation effects under 80% intensity did not significantly differ from those under 100% intensity, $p = .613$. There was a significant main effect of facial expression, $F(1, 21) = 4.509$, $p = . 046$, $\eta^2_p = .177$, with larger facilitation effects for disgusted faces than fearful faces. There was no significant interaction between intensity and facial expression, $F(1.34, 28.15) = 0.23$, $p = .705$.

Table 5.5. Auditory Facilitation Effects on Different levels of Intensity of Facial Expressions.

Facial Expression	Level of Intensity			
	40%	60%	80%	100%
Disgust	39.58	29.38	17.71	20.21
Fear	24.77	11.36	6.82	1.82

5.2.2.3 The Analysis for the Magnitude of Auditory Interference Effects on Different Levels of Intensity of Facial Expressions

The data for auditory interference effects presents in Table 5.6. A 4 x 2 mixed ANOVA with level of intensity (40%, 60%, 80%, and 100%) as a within-subjects factor and with facial expression (disgust and fear) as a between-subjects factor was conducted. A main effect of level of intensity for interference effects was found, $F(1.26, 26.46) = 8.45$, $p = .005$, $\eta^2_p = .287$. Interference effects were smallest under the condition of 40% intensity (-5.03), which was significantly lower than all three other conditions, all *ps* $< .05$. Interference effects observed under 80% intensity (-28.49) significantly differed from those under 60% intensity (-20.54), $p = .047$, but did not significantly differed from those under 100% intensity (-27.28), $p > .10$. There was no significant difference between 60% intensity and 100% intensity, $p > .10$. There was no main effect of facial expressions, $F(1, 21) = 0.08$, $p = .777$ and no interaction, $F\ (3, 63) = 0.60$, $p = .618$.

Table 5.6. Auditory Interference Effects of Different Levels of Intensity of Facial Expressions.

Facial Expression	Level of Intensity			
	40%	60%	80%	100%
Disgust	-6.88	-15.63	-28.13	-25.00
Fear	-3.18	-25.45	-28.86	-29.55

5.2.3 Discussion

The experiment 4 examined the relationship between auditory context effects from vocal expressions and level of intensity of facial expressions. Unlike the results in experiment 3, the current experiment showed that the pattern of auditory consistency effects was the same for disgusted and fearful facial expressions. The magnitude of auditory consistency effects did not reveal any significant differences between any two levels of intensity, which is not consistent with the hypothesis that larger consistency effects would be found at lower level of intensity than at higher level of intensity. The results are in line with the study by Mondloch (2012) who found that the magnitude of context effects did not increase when judging subtle facial expressions (level of 40% intensity) compared to judging unambiguous facial expressions (level of 100% intensity).

Like the analyses done in experiment 3, consistency effects were divided into facilitation effects and interference effects. The findings revealed the magnitude of facilitation effects from emotional sounds was dependent on the level of intensity, with larger facilitation effects on the recognition of facial expressions at lower level of intensity (level of 40% intensity) than on the recognition of facial expressions at higher levels of intensity.

The results are consistent with the prediction that larger context effects under lower level of intensity than under higher level of intensity. The magnitude of interference effects observed on the recognition of both facial expressions showed the opposite trend to that of facilitation effects, which replicated the results obtained in experiment 3. The results showed that larger interference effects were found at the three higher levels of intensity (levels of 60%, 80%, and 100% intensity) than at the lowest level of intensity (level of 40% intensity).

5.3 Chapter General Discussion

Experiment 3 and Experiment 4 sought to explore the relationship between the magnitude of context effects and level of intensity of facial expressions. The current experiments revealed different patterns of consistency effects from visual and auditory contextual information on different levels of intensity. It was found that larger magnitude of visual consistency effects was found at lower levels of intensity of facial expressions, whereas the magnitude of auditory consistency effects did not show any significant differences between any two levels of intensity of facial expressions. The findings might be caused by different types of contextual information. In a previous study (de Gelder & Vroomen, 2000), it has been found that emotional neutral-semantic sentences elicited greater consistency effects on the recognition of more ambiguous mixed facial expressions (e.g., 40% happiness-60% sadness) than on the recognition of unambiguous facial expressions (100% sadness), whereas emotional facial expressions did not elicit larger consistency effects on the recognition of ambiguous emotionally neutral-semantic sentences (e.g., 40% happiness-60%fear) than on the recognition of unambiguous ones (100% fear). More studies should carry out to explore the magnitude of consistency effects from visual and auditory contextual information using different levels of intensity of other facial expressions.

Further analyses showed that the findings of the current experiments showed both facilitation effects and interference effects were dependent on the factor of level of intensity. The magnitude of facilitation effects was larger at lower level of intensity of facial expressions, that is, on the recognition of ambiguous facial expressions, which is consistent with the study by van den Stock et al., (2007). The new finding is that the magnitude of interference effect was less at higher level of intensity of facial expression, which showed the opposite tendency to the magnitude of facilitation effect.

As mentioned above, perceptual similarity between facial expressions were called emotion seeds (Aviezer et al., 2008a, 2008b). When the intensity of a facial expression goes down, the intensity of emotion seeds in a facial expression becomes weaker. For example, a disgusted facial expression shares emotion seeds with an angry one, that is, disgust looks similar to angry. The intensity of emotion seeds which share between a disgusted facial expression and an angry one is reduced as the intensity of the disgusted face descends. In other words, the ability of emotion seeds to cause mis-categorising the disgusted face as angry weakens. It might explain that facilitation effects were larger on the recognition of facial expressions at lower level of intensity than those on the recognition of facial expressions at higher level of intensity. As for interference effects, they were smaller under lower level of intensity of facial expressions than under higher level of intensity ones, it might also because emotion seeds are too weak to be activated by incongruent emotional context.

Carroll and Russell (1996) pointed out that if fewer response options were provided, context effects from situational information on recognition of facial expressions would be eliminated. Righart and de Gelder (2008) have found that when the number of response choices was reduced from 3 to 2, consistency effects for disgusted facial expressions were eliminated. Tanaka-Matsumi, Attivissimo, Nelson and D'urso (1995) suggested that consistency effects might be attributed to the number of response options provided. Interestingly, although it was not the main goals of the current two experiments, consistency effects were still observed for facial expressions of disgust and fear when response alternatives were reduced from six (in experiment 1 and 2) to two (in experiment 3 and 4). The results suggest that consistency effects on recognition of facial expressions are robust no matter how many response options were provided.

In general, recognition of different levels of intensity of facial expressions is influenced by simultaneously presented contextual information. The magnitude of facilitation effects and interference effects depends on level of intensity, with larger facilitation effects on lower level of intensity of facial expressions and larger interference effects on higher level of intensity of facial expressions.

Chapter 6 Context Effects and Attentional Resources

The experiments above demonstrated that context effects are robust. Generally, congruent contextual information improves recognition of facial expressions, while incongruent contextual information impairs recognition of facial expressions to different degrees.

It remained unclear, however, whether the process of such context effects depends on attentional resources. In this chapter, level of perceptual load was manipulated to investigate whether attentional resources are required in the processing of context effects on facial expression recognition. As proposed by load theory (Lavie, 1995; Lavie & Tsal, 1994), in situations of low perceptual load, sufficient attentional resources are left and then are involuntarily allocated to task-irrelevant stimuli, subsequently resulting in distractor effects on target processing. If level of perceptual load were increased, attentional resources would be almost occupied by primary tasks, thus distractor effects from task-irrelevant stimuli would be reduced or disappear. Distractor effects here refer to facilitation effects from irrelevant-congruent contextual stimuli and interference effects from irrelevant-incongruent contextual stimuli on target processing. In other words, if context effects from contextual stimuli were observed under both high and low perceptual load, and were not reduced under high perceptual load condition compared to under low perceptual load condition, it suggests that attention resources would not influence the processing of context effects. If context effects were observed to be decreased or eliminated under high perceptual load condition compared to under low perceptual load condition, it is supposed that context effects elicited by contextual stimuli on recognition of facial expressions would be constrained by attentional resources.

6.1 Experiment 5

The purpose of Experiments 5 was to explore whether the observed visual context effects on recognition of facial expressions were dependent on attentional resources by manipulating level of perceptual load. The way of manipulation of perceptual load will be explained in the procedure section.

6.1.1 Method

6.1.1.1 Participants

Twenty-five new participants between 18 to 37 years of age (12 females and 13 males) were recruited to take part in Experiment 5. All participants were right-handed. Thirteen participants took part in the Disgusted-face block, and 12 participants took part in the Fearful-face block. The data from four participants was excluded from the analyses, two from the Disgusted-face block because the wrong test trials were given to them and two (one per each block) because the participants could not recognise the target facial expression. The wrong test trials were because the experimenter made mistakes on the proportion of the number of go trials and no-go trials. They all had given informed consent and were paid NZD 8 for participation. All had normal/corrected-to-normal vision.

6.1.1.2 Apparatus and Stimuli

The apparatus was the same as that described in chapter 2.

The stimuli were the same as the images in the Experiment 3, including congruent disgusted face-disgusted body images, incongruent disgusted face-angry body images, congruent fearful face-fearful body images, and incongruent fearful face-surprised body images. A coloured shape was placed on the centre of the forehead of each image. There were four combinations of the coloured shapes, consisting of a red circle, a red square, a blue circle and a blue square. The width and height of squares was 0.9 x 0.9 cm, and the diameter of circles was 0.9 cm. Isolated faces served as baseline stimuli and were presented in a separate block in each perceptual load block.

6.1.1.3 Design and Procedure

This experiment used a 2 (facial expression: disgust or fear) x 3 (consistency condition: isolated faces, congruent condition, or incongruent condition) x 2 (perceptual load: low load, or high load) mixed factorial design. The factor of facial expression was a between-subjects factor, and the factors of consistency condition and perceptual load were two within-subjects factors.

Each trial consisted of the following sequence (see Figure 6.1). A fixation cross was displayed for 500 ms. A face-body image or a face image was presented until participants pressed one of the designated keys or disappeared after 1830 ms without any response. Feedback was followed and lasted for 500 ms. Finally, a blank screen was shown for 1000 ms.

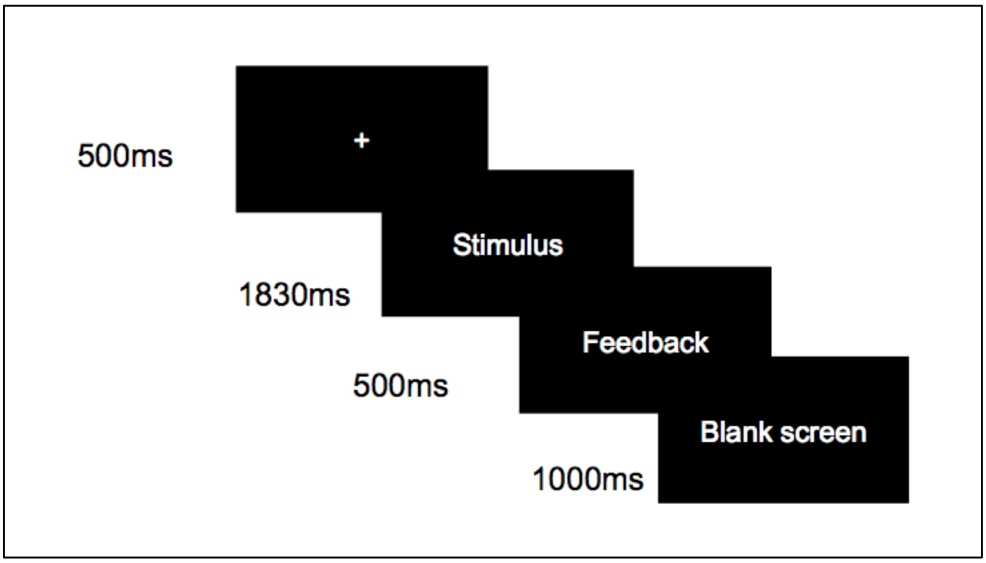

Figure 6.1. An Illustration of a Typical Trial of Experiment 5.

There were two blocks in this test, one is disgusted-face block, and the other is fearful-face block. The experiment used the go/no-go paradigm. In each of the two blocks, participants had to complete the task of facial expression recognition both under low perceptual load and under high perceptual load. The perceptual load was manipulated basing on feature integration theory (Treiman & Gelade, 1980): 1. low load block, participants were instructed to respond to a single feature (colour); 2. high load block, participants were required to respond to the combination of two features (colour and shape). The proportion of go and no-go trials was 3:1 and was the same in both load conditions. In total, there were 144 go trials and 48 no-go trials per each perceptual load block.

Figure 6.2 shows examples of the response cues. For low load trials, the response cue was either the blue colour or the red colour, which was randomly arranged to participants; for high load trials, the response cue was the combination of colour and shape (either a blue square or a red circle), or the reversed combination (either a red square or a blue circle). The order of low and high load conditions was balanced across participants. The task for both load conditions was the same. When the response cue was given to participants, they were required to press ‘c’ for disgusted face and “m’ for angry face in disgusted-face block and ‘c’ for fearful face and ‘m’ for surprised face in fearful-face block by using left and right index finger. They were asked to make a judgment on the emotional faces while ignoring bodily context as accurately and quickly as possible. The response keys representing for the emotions were balanced across participants. When participants were not given the response cue, they were instructed to withhold from pressing any response key. Feedback after each trial was provided to participants on whether they should respond or not, rather than whether the choice of facial expressions recognition was correct or not. “Correct” in blue was on screen when participants pressed one of the response keys after showing with a response cue, or did not press any key after showing with a non-response cue. “Oops” in red was on screen when participants pressed any response key after showing with a non-response cue. “Please faster.” in red on screen when participants did not make a judgment after showing with a response cue. Twenty-four practice trials were given to participants before starting to do the low load block and 48 practice trials before the high load block.

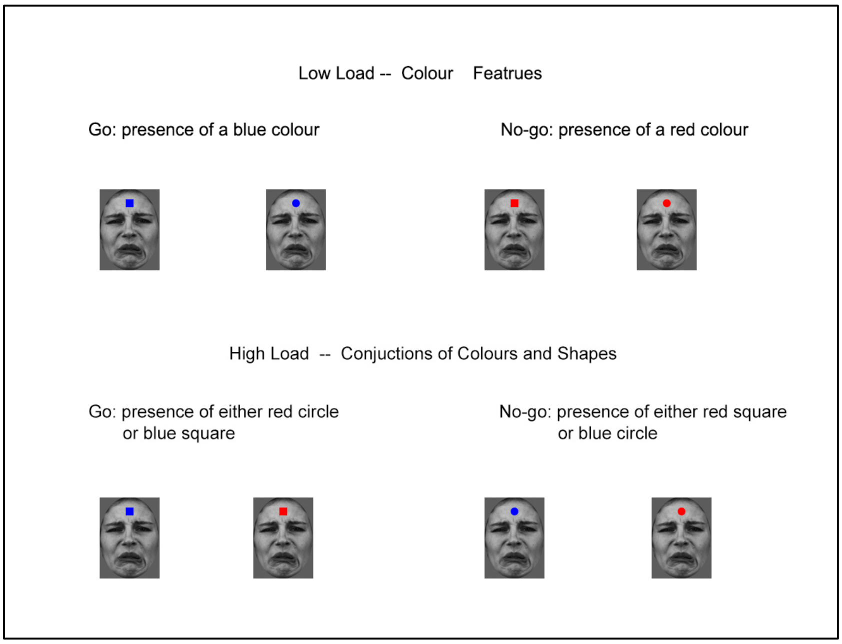

Figure 6.2 Examples of the Response Cues in Go/No-go trials.

6.1.2 Results

Signal Detection Theory (SDT) was adopted to analyse performance on the perceptual load task (the go/no-go task). Then the following analyses were carried out on the data of accuracy of target facial expressions and reaction times (RT) on correct response.

6.1.2.1 Analysis for d' to Assess the Manipulation of Perceptual Load

The sensitivity index, d', was calculated in terms of Z(hits) – Z(false alarms) to ascertain whether the manipulation of perceptual load was effective or not. A hit was defined as that in a go trial, participants pressed one of the two designated keys, no matter the response was correct or not in the task of recognition of facial expressions. And a false alarm was defined as a response made in a no-go trial. The d' is detailed in Table 6.1. A 2 x 3 x 2 mixed factorial ANOVA was conducted for d', with facial expression (Disgust, Fear) as a between-subjects factor and with consistency condition (face-only condition, congruent condition, and incongruent condition) and perceptual load (low load, and high load) as within-subjects factors. A main effect of perceptual load for d' was found, $F(1, 19) = 31.35$, $p < .001$, $\eta^2_p = .623$, with d' = 3.69 under low load larger than d' = 2.83 under high load. A marginal main effect of emotion was found, $F(1, 19) = 3.62$, $p = .072$, $\eta^2_p = .160$, with d'= 3.52 for disgusted faces larger than d' = 3.00. There was no significant main effect of consistency condition for d', $F(2, 38)$ =

2.44, p = .101 . Neither two-way interactions nor three-way interaction was found, all ps > .10.

Table 6.1. d' of Target Facial Expressions with Bodily Expressions under Low Perceptual Load and High Perceptual Load (Visual Context under Visual Load).

Facial Expression	Face-only Condition		Congruent Condition		Incongruent Condition	
	Low Load	High Load	Low Load	High Load	Low Load	High Load
Disgust	3.83 (0.18)	3.14 (0.36)	3.83 (0.17)	3.06 (0.31)	3.91 (0.13)	3.37 (0.17)
Fear	3.36 (0.27)	2.10 (0.20)	3.71 (0.26)	2.62 (0.27)	3.48 (0.24)	2.70 (0.34)

- SEM in Parentheses.

6.1.2.2 Analysis for Accuracy of the Recognition of Facial Expressions

The data of accuracy is presented in Table 6.2. A three-way mixed ANOVA with facial expression (disgust, or fear) as a between-subjects factor and with consistency condition (face-only condition, congruent condition, or incongruent condition) and perceptual load (low load, or high load) as within-subjects factors was conducted, showing a main effect of consistency condition, $F(1.30, 24.75) = 17.86$, $p < .001$, η^2_p

= .484. The accuracy of congruent condition (66.79%) significantly higher than that of isolated faces (49.75%), $p < .001$ and higher than that of incongruent condition (39.11%), $p = .009$, and with the accuracy of incongruent condition significantly less than that of isolated faces, $p < .001$. There were no main effect of perceptual load, $F(1, 19) < 0.001$, $p > .10$, and no main effect of facial expression, $F(1, 19) = 0.001$, $p > .10$. The two-way interaction between facial expression and perceptual load was non-significant, $p > .10$.

Table 6.2. Mean Accuracy of Target Facial Expressions with Bodily Expressions under Perceptual Load (Visual context under Visual Load) (%).

Facial Expressions	Face-only Condition		Congruent Condition		Incongruent Condition	
	Low Load	High Load	Low Load	High Load	Low Load	High Load
Disgust	54.58 (8.03)	55.00 (5.38)	79.79 (9.76)	67.29 (7.43)	15.42 (5.73)	38.75 (5.92)
Fear	46.59 (3.12)	42.80 (4.58)	61.17 (6.14)	58.90 (5.28)	53.79 (5.23)	48.48 (3.91)

- SEM in Parentheses.

The two-way interaction between facial expression and consistency condition was significant, $F(1.30, 24.75) = 9.88$, $p = .002$, $\eta^2_p = .342$. For disgusted faces, the accuracy of congruent condition (73.54%) was significantly greater than that of isolated faces (54.79%), $p = .003$ and greater than that of incongruent condition (27.08%), $p < .001$. The accuracy of incongruent condition significantly less than that of isolated faces, $p < .001$. For fearful faces, the significant difference was only observed between congruent condition (60.04%) and isolated faces (44.70%), $p = .008$. There were no other significant differences between incongruent condition (51.14%) and isolated faces and between congruent condition and incongruent condition, both $ps > .10$.

The two-way interaction between consistency condition and perceptual load was significant, $F(1.94, 36.79) = 5.89$, $p = .006$, $\eta^2_p = .237$. Under low perceptual load, the accuracy of congruent condition (70.48%) was significantly greater than that of incongruent condition (34.60%), $p < .001$ and greater than that of isolated faces (50.59%), $p = .001$. The accuracy of incongruent condition significantly less than that of isolated faces, $p = 0.004$. Under high perceptual load, the accuracy of congruent condition (63.10%) was significantly greater than that of incongruent condition (43.62%), $p = .006$ and greater than that of isolated faces (48.90%), $p = .004$. There was no significant difference between incongruent condition and isolated faces, $p > .10$.

The three-way interaction was significant, $F(1.94, 36.79) = 8.20$, $p = 0.001$, $\eta^2_p = 0.301$. Decomposing the three-way interaction showed that for disgusted facial expressions under low load condition, the accuracy of congruent condition (79.79%) was significantly greater than that of incongruent condition (15.42%), $p < 0.001$ and greater than that of isolated faces (54.58%), $p = .001$. The accuracy of incongruent condition significantly less than that of isolated faces, $p < .001$. For disgusted facial expressions under high load condition, the accuracy of congruent condition (67.29%) was significantly greater than that of incongruent condition (38.75%), $p = .006$, and was marginally different from that of isolated faces, $p = .066$. The accuracy of incongruent condition significantly less than that of isolated faces, $p = .009$.

For fearful facial expressions under low load condition, the accuracy of congruent condition (61.18%) was significantly greater than that of isolated faces (46.59%), $p = .028$. The accuracy of incongruent condition (53.79%) was not significantly different from that of congruent condition and that of isolated faces, both *ps* > .10. For fearful facial expressions under high perceptual load, the accuracy of congruent condition (58.90%) was significantly greater than that of isolated faces (42.80%), $p = .015$. The accuracy of incongruent condition (48.49%) was not significantly different from that of congruent condition and that of isolated faces, both *ps* > .10.

In order to directly compare facilitation effects and interference effects between low load condition and high load condition, a 2 x 2 repeated measures ANOVA with within-subjects factors type of context effect (facilitation effects, interference effects) and perceptual load (low load, and high load) was perform for disgusted facial expressions. The details of facilitation effects and interference effects are listed in

Table 6.3. For disgusted facial expressions, a main effect of type of context effect was not significant, $F(1, 9) = 2.69$, $p = .136$, without significant differences between facilitation effects (18.75) and interference effects (27.71). A main effect of perceptual load was significant, $F(1, 9) = 12.64$, $p = .006$, $\eta^2_p = .584$, with larger context effects under low load condition (32.19) than under high load condition (14.27). The interaction between type of context effect and perceptual load was not significant, $p > .10$.

For fearful facial expressions, there were neither significant interference effects under low load nor significant interference effects under high load. Thus, a paired-sample t test was conducted for facilitation effects. The results showed that there were no significant differences of facilitation effects between low load condition and high load condition, $t(10) = 0.294$, $p > .10$.

Table 6.3. Facilitation and Interference Effects under Perceptual Load (Visual Context under Visual Load).

Facial	Facilitation effects		Interference effects	
Expression	Low Load	High Load	Low Load	High Load
Disgust	25.21 (6.48)	12.29 (3.65)	-39.17 (8.29)	-16.25 (6.80)
Fear	14.58 (6.04)	16.10 (7.60)	7.20 (5.40)	5.68 (4.02)

- SEM in Parentheses

6.1.2.3 Analysis for RTs of the Recognition of Facial Expressions

Mean reaction times (RTs) (listed in Table 6.4) were computed for each participant when making a correct response on facial expression recognition task. A three-way mixed ANOVA with facial expression (disgust or fear) as a between-subjects factor and with consistency condition (face-only condition, congruent condition, or incongruent condition) and perceptual load (low load, or high load) as within-subjects factors was conducted. A main effect of consistency condition was significant, $F(2, 38) = 4.93$, $p = .013$, $\eta^2_p = .206$. RTs for incongruent condition (1007.25 ms) were significantly longer than those for congruent condition (955.03 ms), $p = .012$, and longer than those for isolated faces (947.62 ms), $p = .014$. There was no significant difference of RTs between congruent condition and isolated faces, $p > .10$. A main effect of perceptual load was significant, $F(1, 19) = 10.26$, $p = .005$, $\eta^2_p = .351$, with RTs under low load (900.01 ms) significantly less than RTs under high load (1039.92 ms), which confirmed that the manipulation of perceptual load was effective. There was no main effect of facial expression, $F(1, 19) = 0.675$, $p > .10$. The interaction between perceptual load and facial expression was marginally significant, $F(1, 19) = 3.88$, $p = .064$, $\eta^2_p = .170$. All other two-way interactions and the three-way interaction were non-significant, all *ps* > .10. The results showed that there was no accuracy-speed trade-off.

Table 6.4. Mean Reaction Times (ms) on Correct Responses of Target Facial Expressions with Bodily Expressions under Perceptual Load

Facial Expression	Face-only Condition		Congruent Condition		Incongruent Condition	
	Low Load	High Load	Low Load	High Load	Low Load	High Load
Disgust	871.11 (55.08)	920.11 (58.50)	883.57 (70.66)	953.42 (74.34)	971.80 (65.56)	1014.53 (38.07)
Fear	893.90 (73.16)	1105.36 (92.80)	877.46 (65.21)	1105.68 (69.65)	902.23 (82.08)	1140.43 (77.23)

- SEM in Parentheses.

6.1.3. Discussion

The goal of this experiment was to investigate whether context effects from congruent and incongruent bodily expression to facial expression are dependent on attentional resources. Based on Load Theory (Lavie & Tsal,1994), under low perceptual load, enough attentional resources are left and are involuntarily allocated to the processing of task-irrelevant distractor, whereas under high perceptual load, fewer attentional resources are left to process the task-irrelevant distractor. In other words, if distractor effects (facilitation effects and interference effects) from task-irrelevant stimuli are reduced or eliminated under high load condition compared to low load condition, distractor effects are modulated by attentional resources, whereas if distractor effects are not influenced by perceptual load, distractor effects from task-irrelevant stimuli are free of attentional resources.

The analyses for disgusted facial expressions showed that congruent bodily expressions benefited recognition of facial expressions, with higher accuracy compared to isolated disgusted facial expressions, while incongruent bodily expressions impaired the recognition of facial expression, with lower accuracy and longer RTs compared to isolated disgusted faces. It was found that for disgusted facial expressions, context effects (facilitation effects and interference effects) from task-irrelevant bodily expressions were reduced as level of perceptual load was increased. In the current experiment, under high perceptual load, fewer attention resources are left for processing of task-irrelevant bodily expressions, including congruent and incongruent bodies. The findings suggested that both facilitation effects and interference effects on the recognition of disgusted facial expressions depend on the amount of attentional resources.

For fear facial expressions, recognition of facial expressions was also improved when paired with fearful bodily expressions, with higher accuracy compared to isolated fearful faces. By contrast, the accuracy analysis did not show any significant interference effects from incongruent surprised bodies on the recognition of fearful faces under both perceptual load conditions. Interference effects on fearful facial expression recognition were reflected by the results of RTs that longer RTs were found when fearful faces paired with surprised bodies, in contrast to isolated fearful faces. Unlike disgusted facial expressions, facilitation and interference effects on the recognition of fearful facial expressions were not significantly different from low load condition to high load condition, which suggests that facilitation effects and interference effects on recognition of fearful faces are not constrained by attentional resources. The findings were consistent with an ERP study conducted by Meeren et al. (2005). Their results showed incongruent face-body stimuli enhanced the amplitude of early P1 component than congruent face-body stimuli, which indirectly lends support to the notion the context effects to facial expression did not require attentional resources.

The distinction of processing of context effects between fear and disgust provides new evidence to the dissociated pattern of processing between these two facial expressions. Righart and de Gelder (2008) have found different patterns of context effects on recognition of fearful faces and disgusted faces. In their study, context effects on recognition of fearful faces and disgusted faces exist when completing three-alternative force choice task, whereas when completing two-alternative force choice task, context effects on fearful faces were observed, but no context effects on disgusted facial expressions. Such different patterns for fearful and disgusted facial expressions were found in another study (Lee, Kang, Lee, Namkoong, & An, 2011). Their results indicated a differential priming effect for subliminal fear and disgusted faces, with a stronger subliminal priming effect for fearful faces relative to disgusted faces. The different pattern for processing fearful and disgusted faces might be attributed to relatively separate underlying neural structures for them. Previous studies (Calder et al., 2001; Phillips et al., 1997; Phillips et al., 2004) indicated that the processing of fearful facial expressions mainly activates amygdala, whereas the processing of disgusted facial expressions mainly elicits the activation of insula-basal ganglia system. It would be interesting to explore whether different patterns of context effects would happen to other facial expressions and to further explore the underlying mechanism of different patterns between fearful faces and disgusted one.

6.2 Experiment 6A

Experiment 5 found different patterns of the relationship between visual context effects and attention resources for fearful and disgusted faces. It is found that facilitation effects and interference effects on the recognition of disgusted faces need attentional resources, while facilitation effects and interference effects on the recognition of fearful facial expression are free of attentional resources.

In this experiment, I manipulated level of perceptual load in visual modality (as in Experiment 5) to investigate the relationship between auditory context effects and attentional resources.

6.2.1 Method

6.2.1.1 Participants

Twenty-three new participants between 18 to 36 years of age (13 females and 10 males) were recruited to take part in Experiment 6A. All participants were right-handed except one was left-handed. Twelve participants took part in the Disgusted-face block, and 11 participants took part in the Fearful-face block. One participant was excluded from the analyses, because the participant cannot recognise target facial expression. They all had given informed consent and were paid NZD 8 for participation. All had normal/corrected-to-normal vision and normal hearing ability.

6.2.1.2 Apparatus and Stimuli

The apparatus used was the same as that described in Chapter 2.

The images of disgusted and fearful face images were used in this experiment. The affective bursts of disgust and fear were presented simultaneously with the face images to create emotional-congruent pairs (that is, disgusted face-disgusted sound compounds and fearful face-fearful body compounds) and emotional-incongruent pairs (that is, disgusted face-fearful sound compounds and fearful face-disgusted sound compounds). The response cues in go/no-go task were the same as those in experiment 5.

6.2.1.3 Design and Procedure

The design was the same as that in experiment 5, with facial expression (disgust or fear) as a between-subjects factor, and with consistency condition (face-only condition, congruent condition, or incongruent condition) and perceptual load (low load or high load) as two within-subjects factors.

The procedure was the same as that in Experiment 5 except for the arrangement of response keys. Participants were asked to press ‘c’ or ‘m’ to indicate whether the face was fearful or disgusted, meanwhile ignoring the vocal expressions, when the appropriate response cue appeared.

6.2.2 Results

6.2.2.1 Analysis for d’ to Assess the Manipulation of Perceptual Load

d’ of target facial expression with affective sound is presented in Table 6.5. A mixed factorial ANOVA was also conducted for d’, with facial expression (Disgust, Fear) as one between-subjects factor and with consistency condition (face-only condition, congruent condition, and incongruent condition) and perceptual load (low load, and high load) as within-subjects factors. The results were similar to those in experiment 5. A significant main effect of perceptual load for d' was found, $F(1, 20) = 65.19$, $p < .001$, $\eta^2_p = .765$, with d' = 3.72 under low load larger than d' = 2.63 under high load. There were no significant main effect of consistency condition for d', $F(2, 40) = 1.41$, $p = .257$, and no main effect of facial expression, $F(1, 20) = 0.29$, $p = .598$. Neither two-way interactions nor three-way interaction were found, all *ps* > .10.

Table 6.5. d' of Target Facial Expression with Affective Sound under Low and High Perceptual Load (Auditory Context under Visual Load)

Facial Expression	Face-only Condition		Congruent Condition		Incongruent Condition	
	Low Load	High Load	Low Load	High Load	Low Load	High Load
Disgust	3.55 (0.16)	2.45 (0.24)	3.73 (0.13)	2.78 (0.25)	3.56 (0.15)	2.71 (0.27)
Fear	3.88 (0.12)	2.56 (0.23)	3.84 (0.10)	2.79 (0.20)	3.79 (0.14)	2.52 (0.21)

- SEM in Parentheses

6.2.2.2 Analysis for Accuracy of the Recognition of Facial Expressions

The data of accuracy is presented in Table 6.6. A three-way mixed ANOVA with facial expression (disgust, or fear) as a between-subjects factor and with consistency condition (face-only condition, congruent condition, or incongruent condition) and perceptual load (low load, or high load) as within-subjects factors was conducted. A main effect of consistency condition was significant, $F(1.26, 25.22) = 31.38$, $p < .001$, $\eta^2_p = .611$. The accuracy of congruent condition (74.34%) was significantly higher than that of isolated faces (53.74%), $p < .001$ and higher than that of incongruent condition (31.06%), $p < .001$, and the accuracy of incongruent condition was significantly less than that of isolated faces, $p < .001$. A main effect of perceptual load was not significant, $F(1, 20) = 0.138$, $p > .10$. There was no main effect of facial expression, $F(1, 20) = 0.135$, $p > .10$. All the two-way interactions were not significant, all *ps* > .10. The three-way interaction was non-significant, $p > .10$.

Table 6.6. Mean Accuracy of Target Facial Expressions with Affective Sounds under Low and High Perceptual Load (Auditory Context under Visual Load) (%)

Facial Expression	Face-only Condition		Congruent Condition		Incongruent Condition	
	Low Load	High Load	Low Load	High Load	Low Load	High Load
Disgust	50.19 (3.18)	51.33 (4.35)	81.63 (6.93)	74.81 (4.31)	28.60 (7.28)	35.42 (8.20)
Fear	58.33 (3.84)	55.11 (4.94)	72.35 (7.95)	68.56 (5.72)	30.68 (7.99)	29.55 (3.40)

- SEM in Parentheses

6.2.2.3 Analysis for RTs of the recognition of Facial Expressions

RTs (listed in Table 6.7) were computed for each participant when making a correct response on facial expression recognition task. A three-way mixed ANOVA with facial expression (disgust, or fear) as a between-subjects factor and with consistency condition (face-only condition, congruent condition, or incongruent condition) and perceptual load (low load, or high load) as within-subjects factors was conducted. A main effect of perceptual load was significant, $F(1, 20) = 18.15$, $p < .001$, $\eta^2_p = .476$, with RTs under low load (896.82 ms) significantly less than RTs under high load (1079.93 ms), which confirmed that the manipulation of perceptual load was effective.

There were no main effects of facial expression and consistency condition, both ps > .10. All the three two-way interactions and the three-way interaction were non-significant, all ps > .10. The results showed that there was no accuracy-speed trade-off.

Table 6.7. Mean Reaction times (ms) on correct responses of target facial expressions with affective sound under perceptual load (Auditory context under visual load)

Facial Expression	Face-only Condition		Congruent Condition		Incongruent Condition	
	Low Load	High Load	Low Load	High Load	Low Load	High Load
Disgust	984.63 (44.99)	1103.60 (76.77)	922.15 (44.04)	1079.47 (64.27)	959.59 (73.81)	1097.09 (79.17)
Fear	813.93 (54.27)	1007.99 (55.56)	808.38 (40.25)	1072.55 (68.08)	892.24 (78.82)	1118.87 (78.72)

- SEM in Parentheses

6.2.3 Discussion

Experiment 6A was to find out whether auditory context effects from congruent and incongruent vocal expressions on recognition of facial expressions is dependent on attentional resources. In this experiment, higher accuracy was found for congruent face-sound pairs than for isolated face stimuli under low load and high load conditions. No significant interaction shows that the magnitude of facilitation effects under low load did not significantly differ from that of facilitation effects under high load. The analysis of RTs did not show there was benefit from congruent vocal expressions to facial expression under both load conditions.

The results showed that RTs for incongruent pairs did not significantly differ from those for isolated face stimuli under low load and under high load. The accuracy for incongruent face-sound pairs was lower than that for isolated face stimuli under both low load and high load condition. No significant interaction shows that the magnitude of interference effect was not significant different from low load condition to high load condition. The results indicated that attentional resources might not be needed for context effects, including facilitation and interference effects from vocal expressions on facial expression recognition. The findings are consistent with previous studies (Vroomen et al., 2001), which have found cross-modal interaction and integration proceeds in an automatic way, without the requirement of attentional resources.

6.3 Experiment 6B

The results of Experiment 6A suggest that facilitation effects and interference effects from vocal expressions might not be influenced by attentional resources. However, there is a concern whether attentional resources are centrally limited (Kahneman, 1973) or modality specific (Hancock, Oron-Gilad, & Szalma, 2007; Wickens, 2008). In other words, it is doubtful whether perceptual load in one modality could have an effect on the perception of distractors in another modality. So, in experiment 6B, perceptual load was manipulated in auditory modality to examine whether auditory context effects require attentional resources or not.

6.3.1 Method

6.3.1.1 Participants

Twenty-three new participants between 18 to 40 years of age (18 females and 5 males) were recruited to take part in Experiment 6B. Two participants were left-handed and 21 were right-handed.Twelve participants were arranged in the disgusted-face block, and 11 participants were in the fearful-face block. Two participants (one per each block) were excluded from the analyses, because the participants cannot recognise target facial expression. They all had given informed consent and were paid NZD 8 for participation. All had normal/corrected-to-normal vision and normal hearing ability.

6.3.1.2 Apparatus and Stimuli

The apparatus and stimuli were identical to the details described in Experiment 6A, except for the response cue. The response cue in this experiment was a 2500 Hz pure tone, which was played simultaneously with emotional face images and affective bursts.

6.3.1.3 Design and Procedure

The design and procedure were the same as that in Experiment 6A. The perceptual load was manipulated in auditory modality. For low load trials, participants were asked to respond when hearing a smooth or a trilled tone; for high load trials, participants were asked to respond when hearing a smooth tone from left ear or a trilled tone from right ear, or to respond the reversed combinations (a smooth tone from right ear or a trilled tone from left ear).

6.3.2. Results

6.3.2.1 Analysis for d' to Assess the Manipulation of Perceptual Load

d' of target facial expression with affective sound presents in Table 6.8. A mixed factorial ANOVA was also conducted for d', with facial expression (disgust, or fear) as one between-subjects factor and with consistency condition (face-only condition, congruent condition, or incongruent condition) and perceptual load (low load, or high load) as within-subjects factors. A significant main effect of perceptual load for d' was found, $F(1, 20) = 23.99$, $p < .001$, $\eta^2_p = .657$, with d' = 2.99 under low load larger than d' = 2.09 under high load. There was no significant main effect of consistency condition for d', $F(2, 40) = 0.75$, $p = .478$, and no main effect of facial expression, $F(1, 20) =$

0.002, p = .966. There were no interaction effects between any two factors and no three-way interaction found, all *ps* > .10.

Table 6.8. d' of Target Facial Expressions with Affective Sounds under Low and High Perceptual Load (Auditory Context under Auditory Load).

Facial Expression	Face-only Condition		Congruent Condition		Incongruent Condition	
	Low Load	High Load	Low Load	High Load	Low Load	High Load
Disgust	2.69 (0.24)	2.38 (0.16)	2.69 (0.12)	2.37 (0.21)	3.08 (0.27)	2.11 (0.21)
Fear	3.20 (0.27)	2.25 (0.27)	3.25 (0.14)	1.93 (0.30)	3.06 (0.11)	1.69 (0.22)

- SEM in Parentheses

6.3.2.2 Analysis for Accuracy of the Recognition of Facial Expressions

The data of accuracy is presented in Table 6.9. A three-way mixed ANOVA with facial expression (disgust, or fear) as a between-subjects factor and with consistency condition (face-only condition, congruent condition, or incongruent condition) and perceptual load (low load, or high load) as within-subjects factors was conducted. A main effect of consistency condition was significant, $F(1.30, 24.77) = 26.98$, $p < .001$,

η^2_p = 0.587. The accuracy of congruent condition (69.46%) was significantly higher than that of isolated faces (56.17%), $p < .001$ and higher than that of incongruent condition (32.51%), $p < .001$, and the accuracy of incongruent condition was significantly less than that of isolated faces, $p < .001$. A main effect of load was significant, $F(1, 19) = 11.36$, $p = .003$, $\eta^2_p = .374$, with the accuracy of low load (57.23%) significantly differing from that of high load (48.19%). A main effect of facial expression was significant, $F(1, 19) = 6.58$, $p = .019$, $\eta^2_p = .257$, with the accuracy of disgusted faces (57.55%) significantly higher than that of fearful faces (47.88%). All the two-way interactions were not significant, all *ps* > .10. The three-way interaction was non-significant, $p > .10$.

Table 6.9. Mean Accuracy of Target Facial Expressions with Affective Sounds under Low and High Perceptual Load (Auditory Context under Auditory Load) (%)

Facial Expression	Face-only Condition		Congruent Condition		Incongruent Condition	
	Low Load	High Load	Low Load	High Load	Low Load	High Load
Disgust	69.13 (5.66)	57.20 (5.48)	83.14 (4.21)	69.70 (4.96)	36.74 (7.65)	29.36 (5.97)
Fear	52.08 (6.06)	46.25 (5.57)	69.17 (8.15)	55.83 (5.83)	33.13 (7.33)	30.83 (6.52)

- SEM in Parentheses

6.3.2.3 Analysis for RTs of the Recognition of Facial Expressions

RTs (listed in Table 6.10) were computed for each participant when making a correct response on facial expression recognition task. A three-way mixed ANOVA with facial expression (disgust, or fear) as a between-subjects factor and with consistency condition (face-only condition, congruent condition, or incongruent condition) and perceptual load (low load, or high load) as within-subjects factors was conducted. A main effect of consistency condition was significant, $F(1.47, 27.93) = 9.62$, $p = .002$, $\eta^2_p = .336$. RTs for congruent condition (1125.29 ms) significantly longer than those for isolated faces (1029.59 ms), $p = .005$, but did not significantly differ from those for incongruent condition (1140.69 ms), $p > .10$. RTs for incongruent condition was significantly longer than those for isolated faces, $p = .003$. A main effect of perceptual load was significant, $F(1, 19) = 29.25$, $p < .001$, $\eta^2_p = .606$, with RTs under low load (994.56 ms) significantly less than RTs under high load (1202.48 ms), which confirmed that the manipulation of perceptual load was effective. There was no main effectss of facial expression, $p > .10$. All the three two-way interactions and the three-way interaction were non-significant, all $ps \geq .10$. The results showed that there might be accuracy-speed trade-off.

Table 6.10 Mean Reaction Times (ms) on Correct Response of Target Facial Expressions with Affective Sounds under Perceptual Load (Auditory Context under Auditory Load)

Facial Expression	Face-only Condition		Congruent Condition		Incongruent Condition	
	Low Load	High Load	Low Load	High Load	Low Load	High Load
Disgust	955.99 (59.37)	1141.57 (59.33)	1035.11 (68.00)	1226.02 (63.28)	1093.35 (81.18)	1252.15 (74.67)
Fear	933.18 (68.86)	1087.64 (81.81)	993.7 (54.94)	1246.30 (78.72)	956.02 (64.69)	1261.21 (76.94)

- SEM in Parentheses

6.3.2.4 Analysis for Inversed Efficiency (IE) Scores

In order to take accuracy and RTs together into account, inversed efficiency (IE) scores (Townsend & Ashby, 1978) were used to analyse context effects. IE scores were values of dividing mean RTs by accuracy rates. The smaller IE scores, like RTs, represent better performance. IE scores are presented in Table 6.11.

Because of violation of the assumption of homogeneous variance, analysis was performed on a log transformation of the IE scores. A three-way mixed ANOVA with facial expression (Disgust, Fear) as a between-subjects factor and with consistency condition (face-only condition, congruent condition) and perceptual load (low load, and high load) as within-subjects factors was conducted. A main effect of consistency condition was significant, $F(1.07, 20.41) = 20.46$, $p < .001$, $\eta^2_p = .519$. Participants recognised facial expression under congruent condition (3.22) significantly faster than under isolated faces (3.28), $p = 0.024$, and faster than recognised incongruent condition (3.69), $p < 0.001$. The log IE score under the incongruent condition were significantly slower than that under isolated faces were, $p < .001$. A main effect of perceptual load was significant, $F(1, 19) = 25.68$, $p < .001$, $\eta^2_p = .575$, with log IE score under low load (3.32) less than log IE score under high load (3.47). There was no main effect of facial expression, $p > .10$. All the three two-way interactions and the three-way interaction were non-significant, all *ps* > .10.

Table 6.11 Mean Log Transformation of IE scores on Correct Response of Target Facial Expressions with Affective Sounds under Perceptual Load (Auditory Context under Auditory Load)

Facial Expression	Face-only Condition		Congruent Condition		Incongruent Condition	
	Low Load	High Load	Low Load	High Load	Low Load	High Load
Disgust	3.15 (0.05)	3.31 (0.04)	3.09 (0.05)	3.25 (0.03)	3.56 (0.08)	3.73 (0.12)
Fear	3.26 (0.05)	3.38 (0.06)	3.17 (0.05)	3.35 (0.04)	3.68 (0.21)	3.80 (0.18)

- SEM in Parentheses

6.3.2.5 Analysis for d' for Experiment 6A and 6B

Taking results of both experiment 6A and experiment 6B into account, a four-way mixed designed ANOVA was conducted, with two between-subjects factors of facial expression (disgust or fear) and modality of perceptual load (visual or auditory), and two within-subjects factors of consistency condition (face-only condition, congruent condition, or incongruent condition) and perceptual load (low or high). A main effect of perceptual load was significant, $F(1, 39) = 98.64$, $p < .001$, $\eta^2_p = 0.717$, with d' under low load (3.40) greater than that under high load (2.38). A main effect of modality was significant, $F(1, 39) = 21.64$, $p < .001$, $\eta^2_p = .357$, with d' for visual modality (3.18) larger than that for auditory modality (2.56). The interaction between perceptual load and facial expression was significant, $F(1, 39) = 5.48$, $p = .024$, $\eta^2_p = .123$. For disgusted facial expression, the difference of d' between low load and high load (0.75) was significant, $p < .001$. For fearful facial expressions, the difference of d' between low load and high load (1.21) was also significant, $p < .001$. The other main effects, all the other two-way interactions, all the three-way interactions, and the four-way interaction were all non-significant, all *ps* > .10.

6.3.2.6 Analysis for IE Scores for Experiment 6A and 6B

From the above analyses, there might be speed-accuracy tradeoff for context effects. Besides, due to violation of the assumption of homogeneous variance, the following analysis was conducted for on a log transformation of the IE scores. A four-way mixed designed ANOVA was conducted, with two between-subjects factors of facial expression (disgust, or fear) and modality of perceptual load (visual, or auditory), and two within-subjects factors of consistency condition (face-only condition, congruent condition, or incongruent condition) and perceptual load (low load, or high load). A main effect of consistency condition was significant, $F(1.15, 44.95) = 45.85$, $p < .001$, $\eta^2_p = .540$. Facial expressions were recognised faster when paired with congruent sound (3.18) faster than isolated faces (3.27) and faster than when paired with incongruent sound (3.67), all *ps* < .001. A main effect of perceptual load was significant,

$F(1, 39) = 10.12$, $p = .003$, $\eta^2_p = .206$. Facial expressions under low load (3.33) were recognised faster than those under high load (3.42). The main effects of modality and facial expression were non-significant, both *ps* > .10. The interaction between perceptual load and modality was significant, $F(1, 39) = 4.40$, $p = .042$, $\eta^2_p = .101$. For low perceptual load condition, there was no significant difference of log IE scores of facial expressions between the conditions of visual perceptual load (3.33) and auditory perceptual load (3.32), $p > .10$. For high perceptual load condition, the log IE score of recognition of facial expressions under visual perceptual load (3.36) was marginally less than that under auditory perceptual load (3.47), $p = .076$.

6.3.3 Discussion

Experiment 6B set out to explore whether or not context effects are attention-free by manipulating auditory perceptual load. The results showed that under both low and high load conditions, the accuracy for congruent face-sound pairs was higher than for isolated faces. The accuracy analysis showed that the magnitude of facilitation effects did not differ from low load condition to high load condition, suggesting that attentional resources might not be required. However, the data from RTs analysis revealed that RTs for facial expressions when paired with congruent sounds were longer than those for isolated faces, indicating it might be speed-accuracy trade-off. So, the analysis on IE score was conducted. Facilitation effects and interference effects from congruent and incongruent vocal expressions on recognition of facial expressions were found under both perceptual load conditions, and the magnitude of context effects did not depend on level of perceptual load. The results suggest that the auditory context effects, including facilitation effects and interference effects, on the recognition of disgusted and fearful facial expressions proceed without attentional resources.

The analysis performed on the combination of data of Experiment 6A and 6B showed that it is more difficult for participant performing primary task when perceptual load was manipulated in auditory modality than in visual modality. The comparisons of log IE scores between experiment 6A and 6B revealed that irrelevant vocal expressions processed regardless of whether perceptual load was manipulated in visual modality or in auditory modality. The magnitude of facilitation effects was not influenced by perceptual load, while the magnitude of interference effects was even increased under high auditory perceptual load compared to under high visual perceptual load. Collectively, unlike visual context effects, auditory context effects on recognition of facial expressions were not constrained by attentional resources.

6.4. Chapter General discussion

The last three experiments aimed at examining whether context effects on recognition of facial expressions depend on attentional resources. In Experiment 5, the level of visual perceptual load was manipulated to examine whether attention resources are required for visual context effects on facial expression recognition. In Experiment 6A and 6B, visual perceptual load and auditory perceptual load were manipulated respectively to investigate whether auditory context effects on recognition of facial expressions are dependent on attention resources.

The results of the three experiments revealed that task-irrelevant, emotional bodily expressions and vocal expressions either improve or impair the recognition of facial expressions. It has been found that distractor effects (facilitation effects and interference effects) were modulated by perceptual load (in experiment 5). Specifically, when perceptual load was increased, facilitation effects and interference effects on recognition of disgusted facial expressions were reduced.

It is well documented that irrelevant emotional stimuli influence the ongoing visual task (Bradley et al., 2003; Erthal et al., 2005; Hartikainen, Ogawa, & Knight, 2000; Ishai, Pessoa, Bikle, & Ungerleider 2004; Lane, Chua, & Dolan, 1999; Lang et al., 1998; Mourão-Miranda et al., 2003; Okon-Singer, Tzelgov, & Henik, 2007; Pessoa, Kastner, & Ungerlider, 2002; Pessoa, et al., 2002a; Silver et al., 2007; Simpson et al., 2000; Tipples & Sharma, 2000). In these studies, irrelevant emotional stimuli were emotional face stimuli or emotional non-face stimuli, and the ongoing visual tasks were usually unrelated to emotion. To some extent, the current results could also be considered as evidence that irrelevant emotional stimuli affect the processing of primary visual tasks, although the primary visual task in my study refer to tasks of recognition of facial expressions. The findings that distractor effects from congruent and incongruent bodily expressions were reduced under high perceptual load are consistent with the results found by Erthal et al. (2005). Erthal et al. (2005) found that the influence of irrelevant emotional stimuli on the primary task disappeared when attentional resources were more consumed by increasing task load or by increasing blood alcohol concentration. In this study, participants were asked to judge whether the orientation of two bars were the same or not while ignoring the task-irrelevant, emotional stimuli. When the bar orientation task was difficult enough or when participants were under alcohol intoxication while performing the bar orientation task, the task-irrelevant, unpleasant stimuli (images of mutilated bodies) did not slow RTs of the primary task compared to the the task-irrelevant, neutral stimuli. Similar results were also found in the behavioural study of Okon-Singer et al. (2007). The task in this study was to indicate the target letter was X or N. The level of perceptual load was manipulated by altering set size from one to six letters. Their results showed that interference effects from irrelevant emotional distractor (i.e., violent situations, mutilated bodies, happy babies) compared to no-picture condition were observed under the easier search conditions, whereas disappeared under hard search conditions.

Unlike visual context effects, the results of the last two experiments showed that facilitation effects and interference effects from vocal expressions were not influenced when perceptual load was increased either in visual modality or in auditory modality. The findings suggest that auditory context effects from vocal expressions on categorisation of facial expressions may be unconstrained by attentional resources.

The findings are inconsistent with a cited study (de Jong et al., 2010) that has demonstrated that modality-specific attention is required for visual context effects from facial expressions on perception of auditory emotional voice. The different results on the issue of whether cross-modal interaction requires attentional resources might be caused by: a) different facial expressions were tested. In my study, fearful and disgusted facial expressions were used as stimuli, whereas in de Jong et al.'s (2010) study, two sets of facial expressions were used. One is happiness/sadness, and the other is happiness/fear; b) In my study, I examined whether auditory context effects on the recognition of facial expressions needs attentional resources, whereas de Jong et al. (2010) investigated whether context effects from facial expressions on the recognition of auditory emotional tones requires attentional resources. It has been found that auditory cues affected the localization of visual stimuli, while visual cues could not influence the localization of auditory stimuli. Another study (de Gelder et al., 2005) found the influence of visual stimuli on auditory stimuli differed from the influence of auditory stimuli on visual stimuli. They investigated multisensory integration of emotional information in schizophrenic patients. They found that a decreased effect of emotional voice on recognition of facial expression in patient group compared to normal group, whereas the situation was reversed that an exaggerated effect of facial expression on recognition of emotional voice in patient group compared with normal group. The results suggested that stronger visual than auditory dominance exists in the processing of audiovisual integration. Similar results were also found in another behavioural study (Collignon et al., 2008). The contradictory results were found that the auditory signal (voice dialogues) dominated the visual signal (point-light displays) in emotion perception. (Piwek, Pollick, & Petrini, 2015)

The current results that auditory context effects on facial expression recognition can benefit for our daily life from the perspective of ecology. We can effectively extract information from different modalities, thus make more accurate judgments. The results add new evidence that processing of audiovisual interaction (i.e., McGurk effect, ventriloquist effect) does not depend on attentional resources.

Chapter 7 General Discussion

This study was to investigate the issue regarding context effects on the recognition of facial expressions. Specifically, the purpose of Chapter 4 was to explore whether the recognition of facial expressions is influenced by accompanied context. If so, whether the pattern of context effects is in line with the hypothesis of emotion seeds view. That is, whether the magnitude of context effects relies on level of perceptual similarity between emotions conveyed by a facial expression and a simultaneously presented context (i.e. a bodily expression, a vocal expression). The following Chapter 5 investigated the observed context effects are influenced by the levels of intensity of facial expressions. Finally, Chapter 6 was to examine whether attentional resources are required for the observed context effects on facial expression recognition.

7.1. The Influence of Contextual Stimuli on Recognition of Facial Expressions

This study has demonstrated that both bodily expressions and vocal expressions influence the recognition of facial expressions. The results of experiment 1 and 2 (in chapter 4) indicated that the magnitude of context effects could not be fully predicted by levels of perceptual similarity, which partially lend support to the hypothesis proposed by emotion seeds view. Table 7.1 lists summary of findings of experiment 1 and 2. The possibilities for the current study not finding the pattern of context effects as predicted by emotion seeds view will be discussed in the following sections.

Table 7.1. Summaries of Findings in Experiment 1 and 2

Analyses		Whether the results are consistent with the prediction of emotion seeds view					
		Facial Expressions					
		Anger	Disgust	Fear	Happiness	Sadness	Surprise
Bodily Expressions	Accuracy Analysis	Yes	Yes	Yes	No	Yes	Yes
	Error Analysis 1	Yes	No	Yes	No	Yes	Yes
	Error Analysis 2	Yes	Yes	No	No	Yes	Yes
	Error Analysis 3	Yes	No	Yes	No	Yes	Yes
Vocal Expressions	Accuracy Analysis	Yes	No	No	Yes	Yes	No
	Error Analysis 1	Yes	Yes	Yes	Yes	Yes	Yes
	Error Analysis 2	Yes	Yes	No	Yes	Yes	Yes
	Error Analysis 3	Yes	No	No	Yes	No	Yes

Note: Error analysis 1: to examine whether percentage of mis-categorisation of facial expressions as emotions displayed by high similar context was higher than any other emotions when facial expressions were paired with high similar context; Error analysis 2: to examine whether percentage of mis-categorisation of facial expressions as emotions displayed by high similar context was higher when paired with high similar context than isolated facial expression and when paired with low similar context; Error analysis 3: to examine whether percentage of mis-categorisation of facial expressions when paired with high similar context as emotions displayed by high similar context was higher than percentage of mis-categorisation of facial expressions when paired with low similar context as emotions displayed by low similar context.

Firstly, one possibility was two distinctions between the current study and Aviezer et al's studies (2008b, 2011), for one difference being context stimuli adopted between the current experiments and Aviezer and coworkers' studies (2008b, 2011) and the other difference being experiment design. In the current study, the context stimuli used were pure upper bodies in Experiment 1 and affective non-verbal sound in Experiment 2. Although bodily expressions were evaluated to be reliable emotional context cues in Experiment A, it has been found that it is more difficult to detect an emotion in a bodily expression than in a facial expression (van de Riet & de Gelder, 2008). van de Riet and de Gelder (2008) indicated that a certain emotion could be expressed by multiple bodily expressions, thus emotional body language might not be one-to-one related to specific emotions. In contrast, the context stimuli in Aviezer et al.'s study (2008b, 2011) contained upper bodily expressions with paraphernalia, such as a dirty diaper held by a person, or a tombstone with a face-body compound. A dirty diaper elicits an emotion of disgust, and a tombstone makes us feel sad. It might be the paraphernalia or the scene helped bodily expressions in affecting recognition of facial expressions.

Another difference between the current study and Aviezer et al.'s (2008b) was experiment design. In the current study, within-subjects design was adopted in experiment 1 and 2 to investigate whether emotion seeds view is applicable to all six facial expressions, whereas in Aviezer et al's (2008b) study, only disgusted facial expressions were used as face stimuli to test the hypothesis of emotion seeds views. A study (Yik, Widen, & Russell, 2013) has found that the within-subjects design adopted in studies on recognition of facial expressions led to the judgment of a disgusted facial expression to be affected by a previously presented facial expression. In their study, there were three types of anchor conditions, anger, sadness, and sickness. Participants were randomly assigned to one of them. Each participant was presented with nine trials, with the same order: X, Happiness, Fear, Surprise, X, Fear, Happiness, X, and Disgust, with X replaced with one of the three anchor conditions. They were asked to judge facial expressions on a seven-alternative forced choice rating scale. The results showed that disgusted faces were categorised as disgusted under the angry anchor condition, whereas disgusted facial expressions were categorised as angry under the sad and sick anchor conditions. Therefore, the within-subjects design in the current study might elicit the pattern of context effects differing from that of context effects predicted by emotion seeds view.

Secondly, circumplex (or Dimensional) model might partially explain the results of the current study. According to Circumplex Model, as mentioned in the introduction, emotional facial expressions are recognised on a conjunction of two dimensions of valence (pleasant vs. unpleasant) and arousal (arousal vs. sleepiness) in a circular space (see Figure 1.1 in Chapter 1) (Russell, 1980; Russell & Bullock, 1985). For example, disgusted facial expressions located in bottom-left quadrant are recognised negative in valence and low in arousal. Circumplex Model indicated that facial expressions located in the same quadrant are more similar than facial expressions located in different quadrants. For instance, disgust and sadness located in the same quadrant between 6-9 o'clock are both negative in valence and low in arousal, while disgust and anger located in different quadrants are both negative in valence, but not the same in arousal. Anger is high in arousal, whereas disgust is low in arousal. In terms of Circumplex Model, disgusted facial expressions are more similar to sad facial expressions than to angry facial expressions. In contrast, in terms of emotion seeds view, disgusted faces share more perceptual similarities with angry faces than with sad faces. It is obvious that the similarity between facial expressions predicted by circumplex model is distinct from that predicted by emotion seeds view. In other words, facial expressions that are similar in aspects of valence and arousal might not share emotion seeds (perceptual similarity) between each other. It might thus explain why incongruent angry sound did not elicit significant interference effects on recognition of disgusted facial expressions. However, circumplex model fails to explain why fearful sound elicited larger interference effects to disgusted faces compared to angry sound did. It is supposed to be that fear which is located in the same quadrant as anger should elicit the similar magnitude of context effects to anger should do. Previous study (Mondloch et al., 2013b) explored whether the magnitude of context effects on recognition of fearful, angry, and sad faces varied as a function of perceptual similarity between facial expressions. The similarity is based on either dimensional model or emotion seeds view. Their findings (Mondloch et al., 2013b) revealed that the pattern of context effects on recognition of facial expressions was neither fully consistent with the hypothesis of emotion seeds view nor fully consistent with the hypothesis of circumplex model. In terms of current evidence collected, in order to build a model to predict the relationship of context effect and facial expression, we had better find out how we recognise facial expressions.

Thirdly, emotion seeds might not be sufficient for inducing the pattern of mis-categorisation of facial expressions as predicted by emotion seeds view. Emotion seeds are defined as physical features shared between facial expressions, which keep dormant in isolated faces. According to emotion seeds view, when a facial expression is paired with appropriated context, emotion seeds are activated to influence the recognition of facial expression. In other words, emotion seeds (physical features) turned to be determinants for the recognition of facial expressions. In Aviezer et al.'s (2008b) study, they found angry bodies caused greater interference effects on recognition of disgusted face and disgusted faces were more likely to be mis-categorised as angry when they were paired with angry bodies. However, in the current study, disgusted faces paired with angry bodies were more likely to be mis-categorised as sad than as angry. It might because the activated emotion seeds were not sufficient for us to categorise a disgust face when paired with angry context to be an angry face. This finding seems to replicate the result of a recent study (Bombari et al., 2013) which revealed that individual feature might not be sufficient for recognition of angry facial expressions.

7.2. No Facilitation Effects on the Recognition of Happy and Sad Facial Expressions

The results of Experiment 1 and 2 revealed that the accuracy of congruent face-context pairs was higher than that of isolated faces. However, in both experiments, facilitation effects were not found for recognition of happy and sad faces, which was also found in a study by Boucher (2014). It was reported that happy and sad facial expressions are early developed to transmit reliable signal, while fear/surprise and disgust/anger are lately developed and are then discriminated from each other (Jack & Garrod, 2014). The findings suggested that we could clearly discriminate happiness and sadness from other emotions early in our life, which might explain why no improvement from congruent happy face-context pairs and sad face-context pairs, compared to isolated happy and sad faces.

7.3. Happy Face Advantage

Happy faces are recognised faster and better than other facial expressions (Ambadar, 1994; Feyereisen et al., 1986; Harwood et al., 1999; Kirita & Endo, 1995; Kirouac & Doré, 1983), which are called happy face advantage. Previous studies (Beaudry, Roy-Charland, Perron, Cormier, & Tapp, 2014; Calvo, Nummenmaa, & Avero, 2010; Derntl, Seidel, Kainz, & Carbon, 2009) have found that recognition of happy faces was not even influenced when presented upside-down. The present experiments found that happy facial expressions were recognised better than all the other five facial expressions.

7.4. Perception of Subtle Facial Expressions

It has been found that facial expressions could be recognised at lower levels of intensity (Bartneck & Reichenbach, 2005; Hess et al., 1997; Hoffmann,Kessler, Eppel, Rukavina, & Traue, 2010), which is replicated by the results in Experiment 3 and 4. Subtle facial expressions were recognised better when presented with congruent bodily and vocal expressions, while recognition of subtle facial expressions was impaired when presented with incongruent bodily and vocal expressions.

The main purpose of these two experiments was to investigate the relationship between the magnitude of consistency effects and different levels of intensity of facial expressions (Type B ambiguous facial expressions). As mentioned in the introduction, the magnitude of consistency effects depends on the level of ambiguity of emotional facial expressions (Trope, 1986). Previous studies (van den Stock et al., 2007; de Gelder & Vroomen, 2000) have already demonstrated that the magnitude of consistency effects was larger for Type A of ambiguous facial expressions (e.g. blends of angry and disgusted facial expressions) than for unambiguous facial expressions (e.g. angry or disgusted facial expressions). Consistency effects in previous studies refer to differences of the recognition of facial expressions between congruent condition and incongruent condition. The experiment 3 found the similar results as previous studies that consistency effects from bodily expressions were larger at lower levels of intensity of facial expressions, whereas the experiment 4 did not replicate the results of the experiment 3, revealing that the magnitude of consistency effects from vocal expressions was not influenced by levels of intensity.

Further, consistency effects were decomposed into facilitation effects and interference effects. To my knowledge, this is the first study to separately investigate whether facilitation effects and interference effects were in line with the hypothesis that larger effects occur for lower levels of intensity of facial expressions (that is, an ambiguous facial expression). The results were consistent with the hypothesis, showing that larger magnitude of facilitation effects at lower levels of intensity of facial expressions. A possibility is that participants were biased to make a response on bodily expressions when the level of intensity was decreased. It seems that if subtle facial expressions could not provide reliable cues as unambiguous facial expressions, they would make judgments on bodily expressions although they were asked to categorise facial expressions. Based on this explanation, larger magnitude of interference effects would be expected to occur at lower level of intensity of facial expressions, just as the pattern of the magnitude of facilitation effects. However, it fails to explain the pattern of the magnitude of interference effects. The current results showed the opposite trend of interference effects to facilitation effects, with smaller magnitude of interference effects at lower level of intensity of facial expressions.

Another possibility for the opposite trend of facilitation effects and interference effects is the characteristic of Type B of ambiguous facial expressions. As mentioned before, Type B ambiguous facial expressions is created from neutral faces to unambiguous facial expressions. Reducing level of intensity of facial expressions results in increasing level of ambiguity of facial expressions, thus it is more difficult to recognise lower levels of intensity of facial expressions compared to higher levels of intensity of facial expressions (Bartneck & Reichenbach, 2005; Hess et al., 1997). When level of intensity of a facial expression was decreased, perceptual similarity between facial expressions was reduced as well. That is, the intensity of emotion seeds in a facial expression becomes weaker as the intensity of a facial expression decreased. Therefore, the reason that facilitation effects are larger under lower level of intensity of facial expressions compared to higher intensity ones might be the ability of emotion seeds to give rise to mis-categorising a facial expression as another one weakens. It could also explain smaller interference effects under low intensity than under high intensity, because emotion seeds are too weak to be activated by incongruent context.

7.5. Context Effects and Attentional Resources

Until now, it is unresolved whether emotion processing is prioritized and is independent of attentional resources. As mentioned in the introduction, some studies (Bradley et al., 1996; Buodo et al., 2002; Pereira et al., 2006; Schupp et al., 2003) suggested that emotional perception is automatic and without restriction of attentional resources, while other studies (De Cesarei et al.,2009; Erthal et al., 2005; Lim & Pessoa, 2008; Mitchell et al.,2007; Pessoa et al., 2002a; Pessoa et al., 2005; Schupp et al., 2007; Silver et al.,

2007) argued for that emotion perception is modulated by attentional resources. When it comes to whether context effects need attentional resources or not, the results are also mixed, with some studies supporting that context effects are free of attention (Aviezer et al., 2011; Meeren et al., 2006; Righart & de Gelder, 2008; Vroomen et al., 2001c), other studies (de Jong et al., 2010; Gu et al., 2013) supporting context effects require attentional resources.

Experiment 5, 6A and 6B were conducted to examine whether context effects are free of attentional resources. These experiments made use of the concept of perceptual load (Lavie, 1995) to explore this issue. It was proposed that when perceptual load is low, enough attention resources are left and are available to the processing of task-irrelevant stimuli. As perceptual load is increased, fewer attentional resources will be left, thus the processing of task-irrelevant stimuli will be decreased or eliminated. In these experiments, task-irrelevant stimuli were bodily expressions in Experiment 5 and vocal expressions in Experiment 6A and 6B.

The results in experiment 5 showed that the magnitude of visual facilitation effects and interference effects for the recognition of facial expressions was decreased as level of perceptual load was increased. The findings suggested that visual context effects require attentional resources, at least for visual context effects on the recognition of disgusted facial expressions. In contrast, the results in experiment 6A and 6B revealed that facilitation effects and interference effects from vocal expressions are not constrained by attentional resources.

The findings that attentional resources required for visual context effects (facilitation effects and interference effects) do not contradict the neural evidence obtained by ERP studies which explored context effects from visual and auditory contextual stimuli on facial expression recognition only under no load condition. Under no load condition, previous ERP studies (Meeren et a., 2005; Pourtois et al., 2000; Righart & de Gelder, 2008) suggested that context effects from bodily expressions (Meeren et a., 2005) and from vocal expressions (Righart & de Gelder, 2008; Pourtois et al., 2000) on recognition of facial expressions happen at the early processing stage. How about the situation under high perceptual load? To date, there is no ERP study conducted to explore the issue. Moreover, Luck and Hillyard (1999) proposed that attentional resources could operate at different processing stages. It has been reported that not only can attentional resources modulate late processing stages (such as, working memory) (Driver & Frith, 2000; Driver & Spence, 2000; Luck & Hillyard, 1999), but they can also have an influence at early stages of processing, even early at sensory processing (Luck & Hillyard, 1999). The findings that attentional resources modulate early stages of processing provide evidence consistent with the current results.

Additionally, previous fMRI studies might also be in line with the results that context effects require attentional resources. Mobbs et al. (2006) used fMRI to investigate the neural structure concerning visual contextual influences on the rating of facial expressions. The contextual images were presented by “zooming in or out of the image” to produce “a dynamic ‘movie’ effect”. Each contextual image was presented for 4 s. After a jittered inter-stimulus interval (ISI) varying between 4s, 6s, 8s, a face with a neutral or subtle facial expression was presented for 750ms. At last, participants were asked to rate emotional facial expressions on a two-dimensional rating scale basing on Russell’s circumplex model. fMRI data revealed that faces paired with emotional contextual movies increased BOLD responses in the bilateral temporal pole, anterior cingulate cortices (ACC), amygdala, and bilateral superior temporal sulcus (STS) compared to identical faces paired with neutral context movies. The results were indirectly replicated in another study (Righart & de Gelder, 2008a) where the results showed that N170 amplitudes went more negative when fearful faces were in congruent (fearful) scenes than in incongruent (happy) scenes. It was suggested that N170 amplitudes are related with BOLD responses in the STS (Hadjikhani, Kveraga, Naik, & Ahlfors, 2009; Henson et al., 2003; Itier & Taylor, 2004; Nguyen & Cunnington, 2014; Shibata et al., 2002). Superior temporal sulcus (STS) is a region containing face-selective cells (Tsao, Freiwald, Tootell, & Livingstone, 2006) and is associated with the perception of changeable aspects of the face, such as emotional expression, lip reading and eye direction (Haxby, Hoffman, & Gobbini, 2000). Redcay (2008) has reported that the STS is activated by stimuli conveyed social communicative significance, such as body movements, vocal sound, and head orientation. It has been found that the posterior STS plays a crucial role in audiovisual integration (Campanella

& Belin, 2007; Beauchamp, 2005). The STS is connected with both parietal and frontal systems that are involved in discriminative functioning and the allocation of attention (Hein & Knight, 2008). Indeed, it has been found that the activity of STS region was enhanced in the allocation of spatial attention in response to visual cues (Hopfinger, Buonocore, and Mangun, 2000). Besides the STS, the activation of ACC correlates with attentional control, with greater activation of ACC when selectively attending to a stimulus or inhibiting a response to a certain stimulus (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Bush, Luu, & Posner, 2000; Crottaz-Herbette & Menon, 2006; Davis, Hutchison, Lozano, Tasker, & Dostrovsky, 2000; Gehring & Knight, 2000; Carter, Botvinick, & Cohen, 1999; Milham et al., 2001; Paus, 2001; Weissman, Gopalakrishnan, Hazlett, & Woldorff, 2005). These neural findings lend support to the current results that visual context effects on recognition of facial expressions are modulated by attentional resources.

Converging evidence suggests that amygdala plays a crucial role in detecting and responding to threatening information. fMRI studies have revealed that the greater activation of amygdala when presenting fearful facial expressions to participants compared to neutral faces (Adolphs, Russell, & Tranel, 1999; Breiter et al., 1996; Morris et al., 1998). It is commonly suggested that fear is processed fast through the subcortical way, and without requirement of attention resources, when the activation of amygdala was found in completing a task concerning fear. However, a recent study has found that a patient with complete bilateral amygdala lesions could rapidly detect fearful faces among distractor stimuli and non-consciously process those same fearful faces, whereas this individual cannot recognise fearful facial expressions (Tsuchiya, Moradi, Felsen, Yamazaki, & Adolphs, 2009). This result suggested that amygdala does not involve in early processing stages of fearful faces, instead affects the recognition and social judgment. There were two more studies (Bach, Talmi, Hurlemann, Patin, & Dolan, 2011; Piech et al., 2010) lending support to and extending this result. Piech et al. (2010) found that quicker responses for fear-related stimuli (e.g., spider) compared to neutral stimuli were observed for both unilateral amygdala lesions patients and control participants. However, it was argued that the ability to efficiently detect fearful-related stimuli might be due to the intact contralateral amygdala. A more recent study (Bach et al., 2011) revealed that when completing tasks in the attentional blink paradigm, two patients with selective bilateral amygdala lesions were benefitted in recall facilitation for aversive words. The recall facilitation for the two patients was similar to that for control groups. Taken together, the amygdala is not a specialized neural structure for the rapid processing of emotional information.

Besides, it has been doubted that whether there is a straight subcortical pathway to the amygdala (Pessoa & Adolphs, 2011). Pessoa and Adolphs (2010) have proposed a revised function of the amygdala that serves to allocate limited processing resources to stimuli that are most relevant to tasks. Therefore, the activation of amygdala could not evidence that context effects on recognition of facial expressions are free of attentional resources.

7.6. Limitations and Future Directions

7.6.1. Limitations

A number of limitations might be worth mentioning. One limitation of the present study is still images of facial expressions which may lack of ecological validity. Future studies could adopt dynamic displays of facial expressions that are recognised better than still images of facial expressions (Ambadar et al., 2005) to investigate whether the recognition of dynamic facial expressions would be influence by contextual information and whether the pattern of context effects on still facial expression recognition would be similar to that on dynamic facial expressions recognition. Another limitation is two-alternative forced choice task was adopted when investigating context effects on different levels of intensity of facial expressions. Although the current study has reached its aims, future studies could provide more options to further explore the specific pattern of context effects on recognition of subtle facial expressions. For example, which emotion category would be more likely to be chosen when subtle facial expressions were paired with incongruent contextual information? Additionally, future studies could use different levels of intensity of happy, angry, sad, and surprised facial expressions to examine whether the similar pattern of context effects on recognition of different levels of intensity of disgusted and fearful facial expressions would be observed. The last limitation is a limited number of stimuli used in the current study. Future studies could not only utilize more stimuli but also adopt a different set of stimuli to explore the pattern of context effects on recognition of facial expressions.

7.6.2. Future Directions

The current study has found that recognition of facial expressions is influenced by concurrently presented context. However, emotion seeds view can only partially explain the pattern of context effects on facial expression recognition. As discussed earlier, the similarity between facial expressions based on physical characteristics (as used in emotion seeds view) is different from that based on circumplex model. One way to find out how contextual information influences recognition of facial expressions would be that future studies might concentrate on exploring an appropriate way of how to evaluate the similarity between facial expressions. Previous studies have found separate neural systems (Adolphs, 2002; Blair, Morris, Frith, Perrett, & Dolan, 1999; Hennenlotter, 2005; Sprengelmeyer, Rausch, Eysel, & Przuntek, 1998), different adaptive functions (Chapman et al., 2009; Marsh et al., 2005; Öhman & Mineka, 2001; Rozin et al., 1994; Shariff & Tracy, 2011) and different developed time course (Jack, Garrod, & Schyns, 2014) for different facial expressions. Therefore, another direction to explore how contextual information affects facial expression recognition would be to investigate whether there is a consistent pattern of context effects on recognition of all six basic facial expressions or not.

The present results reveal that attentional resources modulate context effects of bodily expressions on recognition of disgusted facial expressions, whereas attentional resources are not required for context effects of bodily expressions on fearful face recognition. Future studies could investigate whether attentional resources would be required for context effects on recognition of the other four facial expressions. Further, more studies might carry on to explore the function of neural structures for emotional processing as such it could deepen and extend our understanding of the relationship between attentional resources and context effects.

7.7. Summary

The study demonstrated that recognition of facial expressions is influenced by simultaneously presented bodily and vocal expressions. The current results showed that emotion seeds view could partially explain the pattern of interference effects on facial expression recognition. At present, the conclusion is that congruent bodily and vocal expressions enhance the performance of facial expression recognition, while incongruent ones impair facial expression recognition. Besides unambiguous facial expressions, recognition of subtle facial expressions is also influenced by accompanied contextual stimuli. The patterns of facilitation effects and interference effects show the opposite trend, with larger facilitation effects at lower levels of intensity of facial expressions and with larger interference effects for facial expressions at higher levels of intensity. Whether attention is required in the processing of context effects was examined under the framework of load theory. The influence from congruent and incongruent bodily expressions on recognition of disgusted facial expressions requires attentional resources, whereas that on fearful facial expressions does not depend on attentional resources. In contrast, context effects from congruent and incongruent vocal expressions on recognition of disgusted and fearful facial expressions are free of attentional resources.

Reference

Adams, R. B., Jr., & Kleck, R.E. (2003). Perceived gaze direction and the processing of facial displays of emotion. *Psychological Science*, *14*(6), 644-647. doi: 10.1046/j.0956-7976.2003.psci_1479.x

Adams, R. B., Jr., & Kleck, R. E. (2005). Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotio, 5*(1), 3-11. doi: 10.1037/1528-3542.5.1.3

Adams,R.B., Jr. Ambaby, N., Marcrae, C.N., & Kleck, R.E. (2006). Emotional expressions forecast approach-avoidance behavior. *Motivation and emotion*, 30(2), 177-186. doi: 10.1007/s11031-006-9020-2

Adolphs, R., Russell, J. A., & Tranel, D. (1999). A role for the human amygdala in recognizing emotional arousal from unpleasant stimuli. Psychological Science, 10(2), 167-171. doi: 10.1111/1467-9280.00126

Adolphs, R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behavioral and cognitive neuroscience reviews*, *1*(1), 21-62. doi: 10.1177/1534582302001001003

Ambadar, Z., Schooler, J. W., & Cohn, J. F. (2005). Deciphering the enigmatic face the importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science*, *16*(5), 403-410. doi: 10.1111/j.0956-7976.2005.01548.x

Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalmic and Physiological Optics, 34*(5), 502-508. doi: 10.1111/opo.12131

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception-London, 33*(6), 717-746. doi:10.1068/p5096

Aviezer, H., Bentin, S., Dudarev, V., & Hassin, R. R. (2011). The automaticity of emotional face-context integration. *Emotion, 11*(6), 1406-1414. doi: 10.1037/a0023578

Aviezer, H., Hassin, R., Bentin, S., & Trope, Y. (2008a). Putting facial expressions back in context. In N. Ambady & J. J. Skowronski (Eds.), *First impressions* (pp. 255-286). New York: Guilford Press. Retrieve from http://cel.huji.ac.il/publications/pdfs/Aviezer_et_al_2008_Chapter_in_First_Impressions.pdf

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., & Bentin, S. (2008b). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science, 19*(7), 724-732.

Aviezer, H., Trope, Y., & Todorov, A. (2012). Holistic person processing: faces with bodies tell the whole story. *Journal of personality and social psychology, 103(1),* 20-37. doi: 10.1037/a0027411.

Bach, D. R., Talmi, D., Hurlemann, R., Patin, A., & Dolan, R. J. (2011). Automatic relevance detection in the absence of a functional amygdala. *Neuropsychologia*, *49*(5), 1302-1305. doi: 10.1016/j.neuropsychologia.2011.02.032

Ballihi, L., Lablack, A., Amor, B. B., Bilasco, I. M., & Daoudi, M. (2015). Positive/Negative Emotion Detection from RGB-D upper Body Images. In *Face and Facial Expression Recognition from Real World Videos* (pp. 109-120). Springer International Publishing. doi: 10.13140/2.1.1796.2246

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of personality and social psychology*, *70*(3), 614-636. doi: 10.1037/0022-3514.70.3.614

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*,*5*(8), 617-629. doi: 10.1038/nrn1476

Bartholow, B. D., Pearson, M. A., Dickter, C. L., Sher, K. J., Fabiani, M., & Gratton, G. (2005). Strategic control and medial frontal negativity: Beyond errors and response conflict. *Psychophysiology, 42*(1), 33-42. doi: 10.1111/j.1469-8986.2005.00258.x

Bartneck, C., & Reichenbach, J. (2005). Subtle emotional expressions of synthetic characters. *International journal of human-computer studies, 62*(2), 179-192. doi:10.1016/j.ijhcs.2004.11.006

Batty, M., & Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions. *Cognitive Brain Research, 17*(3), 613-620. doi:10.1016/S0926-6410(03)00174-5

Beauchamp, M. S. (2005). Statistical criteria in FMRI studies of multisensory integration. *Neuroinformatics*, *3*(2), 93-113. doi: 10.1385/NI:3:2:093

Beaudry, O., Roy-Charland, A., Perron, M., Cormier, I., & Tapp, R. (2014). Featural processing in recognition of emotional facial expressions. *Cognition & emotion*, *28*(3), 416-432. doi: 10.1080/02699931.2013.833500

Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal Affective Voices: a validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior research methods*, *40*(2), 531-539. doi: 10.3758/BRM.40.2.531

Benson, P. J., & Perrett, D. I. (1991). Synthesising continuous-tone caricatures. *Image and vision computing*, *9*(2), 123-129. doi:10.1016/0262-8856(91)90022-H

Bertelson, P., Vroomen, J., De Gelder, B., & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & psychophysics, 62*(2), 321-332. doi: 10.3758/BF03205552

Biehl, M., Matsumoto, D., Ekman, P., Hearn, V., Heider, K., Kudoh, T., & Ton, V. (1997). Matsumoto and Ekman's Japanese and Caucasian Facial Expressions of Emotion (JACFEE): Reliability data and cross-national differences. *Journal of Nonverbal Behavior*, *21*(1), 3-21. doi: 10.1023/A:1024902500935

Bimler, D., & Kirkland, J. (2001). Categorical perception of facial expressions of emotion: Evidence from multidimensional scaling. *Cognition & Emotion, 15*(5), 633-658. doi: 10.1080/02699930126214

Blair, R. J. R., Morris, J. S., Frith, C. D., Perrett, D. I., & Dolan, R. J. (1999). Dissociable neural responses to facial expressions of sadness and anger. *Brain*, *122*(5), 883-893. doi: *http://dx.doi.org/10.1093/brain/122.5.883*

Blair, R. J. R., Mitchell, D. G. V., Peschardt, K. S., Colledge, E., Leonard, R. A., Shine, J. H., ... & Perrett, D. I. (2004). Reduced sensitivity to others' fearful expressions in psychopathic individuals. *Personality and Individual Differences*, *37*(6), 1111-1122. doi:10.1016/j.paid.2003.10.008

Bombari, D., Schmid, P. C., Schmid Mast, M., Birri, S., Mast, F. W., & Lobmaier, J. S. (2013). Emotion recognition: The role of featural and configural face information. *The Quarterly Journal of Experimental Psychology*, *66*(12), 2426-2442. doi: 10.1080/17470218.2013.789065

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624-652.doi: 10.1037/0033-295x.108.3.624

Boucher, J. D., & Carlson, G. E. (1980). Recognition of facial expression in three cultures. *Journal of Cross-Cultural Psychology, 11*(3), 263-280. doi: 10.1177/0022022180113003

Boucher, P. (2014). *Modulation of Facial Expression Perception by Body Context* University of Waterloo, Waterloo, Ontario, Canada. Retrieved from https://uwspace.uwaterloo.ca/handle/10012/8997

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1996). Picture media and emotion: Effects of a sustained affective context. *Psychophysiology, 33*(6), 662-670. doi: 10.1111/j.1469-8986.1996.tb02362.x

Bradley, M. M., Sabatinelli, D., Lang, P. J., Fitzsimmons, J. R., King, W., & Desai, P. (2003). Activation of the visual cortex in motivated attention. *Behavioral neuroscience*, *117*(2), 369. doi: 10.1037/0735-7044.117.2.369

Breiter, H. C., Etcoff, N. L., Whalen, P. J., Kennedy, W. A., Rauch, S. L., Buckner, R. L., ... & Rosen, B. R. (1996). Response and habituation of the human amygdala during visual processing of facial expression. *Neuron*,*17*(5), 875-887. doi:10.1016/S0896-6273(00)80219-6

Brennan, S.E. (1985). Caricature generator: the dynamic exaggeration of faces by computer. *Leonardo, 18(3),* 170-178. Retrieve from http://www.cs.rutgers.edu/~decarlo/671/readings/brennan_1985.pdf

Broadbent, D. (1958). *Perception and communication*. New York: Pergamon Press

Buodo, G., Sarlo, M., & Palomba, D. (2002). Attentional resources measured by reaction times highlight differences within pleasant and unpleasant, high arousing stimuli. *Motivation and Emotion, 26*(2), 123-138. doi: 10.1023/A:1019886501965

Burnham, B. R., Sabia, M., & Langan, C. (2014). Components of working memory and visual selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 40*(1), 391-403. doi: 10. 1037/a0033753

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in cognitive sciences*, *4*(6), 215-222.

Busse, L., Roberts, K. C., Crist, R. E., Weissman, D. H., & Woldorff, M. G. (2005). The spread of attention across modalities and space in a multisensory object. *Proceedings of the National Academy of Sciences of the United States of America, 102*(51), 18751-18756. doi:10.1016/S1364-6613(00)01483-2

Calder, A. J., Young, A. W., Perrett, D. I., Etcoff, N. L., & Rowland, D. (1996). Categorical perception of morphed facial expressions. *Visual Cognition, 3*(2), 81-118. doi: 10.1080/713756735

Calder, A. J., Lawrence, A. D., & Young, A. W. (2001). Neuropsychology of fear and loathing. *Nature Reviews Neuroscience*, *2*(5), 352-363. doi:10.1028/35072584

Calvo, M. G., Nummenmaa, L., & Avero, P. (2010). Recognition advantage of happy faces in extrafoveal vision: Featural and affective processing. *Visual Cognition*, *18*(9), 1274-1297. doi: 10.1080/13506285.2010.481867

Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in cognitive sciences*, *11*(12), 535-543. doi:10.1016/j.tics.2007.10.001

Carretié, L., Hinojosa, J. A., Martín-Loeches, M., Mercado, F., & Tapia, M. (2004a). Automatic attention to emotional stimuli: neural correlates. *Human brain mapping, 22*(4), 290-299. doi: 10.1002/hbm.20037

Carretié, L., Mercado, F., Hinojosa, J. A., Martí, M., & Sotillo, M. (2004b). Valence-related vigilance biases in anxiety studied through event-related potentials. *Journal of affective disorders, 78*(2), 119-130. doi:10.1016/S0165-0327(02)00242-2

Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of personality and social psychology*, *70*(2), 205. doi: 10.1037/0022-3514.70.2.205

Carter, C. S., Botvinick, M. M., & Cohen, J. D. (1999). The contribution of the anterior cingulate cortex to executive processes in cognition. *Reviews in the Neurosciences*, *10*(1), 49-58.

Castelli, F. (2005). Understanding emotions from standardized facial expressions in autism and normal development. *Autism*, *9*(4), 428-449. doi: 10.1177/1362361305056082

Chapman, H. A., Kim, D. A., Susskind, J. M., & Anderson, A. K. (2009). In bad taste: Evidence for the oral origins of moral disgust. *Science*, *323*(5918), 1222-1226. doi: 10.1126/science.1165565

Chen, S., Tian, Y., Liu, Q., & Meta, D. N. (2011, March). Segment and recognize expression phase by fusion of motion area and neutral divergence features. In *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on* (pp. 330-335). IEEE. doi: 10.1109/FG.2011.5771419

Chen, S., Tian, Y., Liu, Q., & Metaxas, D. N. (2013). Recognizing expressions from face and body gesture by temporal normalized motion and appearance features. *Image and Vision Computing*, *31*(2), 175-185. doi:10.1016/j.imavis.2012.06.014

Christie, I. C., & Friedman, B. H. (2004). Autonomic specificity of discrete emotion and dimensions of affective space: a multivariate approach. *International journal of psychophysiology*, *51*(2), 143-153. doi:10.1016/j.ijpsycho.2003.08.002

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain research, 1242*, 126-135. doi:10.1016/j.brainres.2008.04.023

Crottaz-Herbette, S., & Menon, V. (2006). Where and when the anterior cingulate cortex modulates attentional response: combined fMRI and ERP evidence. *Cognitive Neuroscience, Journal of*, *18*(5), 766-780. doi: 10.1162/jocn.2006.18.5.766

Dailey, M. N., Cottrell, G. W., Padgett, C., & Adolphs, R. (2002). EMPATH: A neural network that categorizes facial expressions. *Journal of cognitive neuroscience, 14*(8), 1158-1173. doi:10.1162/089892902760807177

Darwin, C. (1872). *The expression of the emotions in man and animals*. London: J. Murray.

Davis, H., Mast, T., Yoshie, N., & Zerlin, S. (1966). The slow response of the human cortex to auditory stimuli: recovery process. *Electroencephalography and clinical neurophysiology, 21*(2), 105-113. doi:10.1016/0013-4694(66)90118-0

Davis, K. D., Hutchison, W. D., Lozano, A. M., Tasker, R. R., & Dostrovsky, J. O. (2000). Human anterior cingulate cortex neurons modulated by attention-demanding tasks. *Journal of neurophysiology*, *83*(6), 3575-3577. Retrieve from https://jn.physiology.org/content/jn/83/6/3575.full.pdf

De Cesarei, A., & Codispoti, M. (2006). When does size not matter? Effects of stimulus size on affective modulation. *Psychophysiology, 43*(2), 207-215. doi: 10.1111/j.1469-8986.2006.00392.x

De Cesarei, A., Codispoti, M., & Schupp, H. T. (2009). Peripheral vision and preferential emotion processing. *Neuroreport, 20*(16), 1439-1443. doi: 10.1097/WNR.0b013e3283317d3e

de Fockert, J. W., Rees, G., Frith, C. D., & Lavie, N. (2001). The role of working memory in visual selective attention. *Science, 291*(5509), 1803-1806. doi: 10.1126/science.1056496

de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, *7*(3), 242-249. doi:10.1038/nrn1872 de Gelder, B. (2009). Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3475-3484. doi: 10.1098/rstb.2009.0190

de Gelder, B., Teunisse, J. P., & Benson, P. J. (1997). Categorical perception of facial expressions: Categories and their internal structure. *Cognition & Emotion*,*11*(1), 1-23. doi: 10.1080/026999397380005

de Gelder, B., Böcker, K. B., Tuomainen, J., Hensen, M., & Vroomen, J. (1999). The combined perception of emotion from voice and face: early interaction revealed by human electric brain responses. *Neuroscience letters*, *260*(2), 133-136. doi:10.1016/S0304-3940(98)00963-X

de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition & Emotion*, *14*(3), 289-311. doi: 10.1080/026999300378824

de Gelder, B., & Van den Stock, J. (2011). The bodily expressive action stimulus test (BEAST). Construction and validation of a stimulus basis for measuring perception of whole body expression of emotions. *Frontiers in psychology, 2(8),* 181. http://dx.doi.org/10.3389/fpsyg.2011.00181 de Jong, J. J., Hodiamont, P. P. G., & de Gelder, B. (2010). Modality-specific attention and multisensory integration of emotions in schizophrenia: reduced regulatory effects. *Schizophrenia research*, *122*(1), 136-143. doi:10.1016/j.schres.2010.04.010

Delplanque, S., Lavoie, M. E., Hot, P., Silvert, L., & Sequeira, H. (2004). Modulation of cognitive processing by emotional valence studied through event-related potentials in humans. *Neuroscience letters*, *356*(1), 1-4. doi:10.1016/j.neulet.2003.10.014

Derntl, B., Seidel, E. M., Kainz, E., & Carbon, C. C. (2009). Recognition of emotional expressions is affected by inversion and presentation time. *Perception, 38*(12*),*1849-1862. doi: 10.1068/p6448

Deutsch, J. A., & Deutsch, D. (1963). Attention: some theoretical considerations. *Psychological review*, *70*(1), 80-90. doi: 10.1037/h0039515

Dolan, R. J., Morris, J. S., & de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proceedings of the National Academy of Sciences*, *98*(17), 10006-10010. doi: 10.1073/pnas.171288598

Driver, J. (1996). Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. *Nature*, *381*(6577), 66-68. doi:10.1038/381066a0

Driver, J., & Frith, C. (2000). Shifting baselines in attention research. *Nature Reviews Neuroscience*, *1*(2), 147-148. doi:10.1038/35039083

Driver, J., & Spence, C. (2000). Multisensory perception: beyond modularity and convergence. *Current Biology*, *10*(20), R731-R735. doi:10.1016/S0960-9822(00)00740-5

Ducci, L., Arcuri, L., W/Georgis, T., & Sineshaw, T. (1982). Emotion recognition in Ethiopia: The effect of familiarity with Western culture on accuracy of recognition. Journal of Cross-Cultural Psychology, 13(3), 340–351. doi: 10.1177/0022002182013003005

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological review*, *87*(3), 272-300. doi: http://dx.doi.org.ezproxy.massey.ac.nz/10.1037/0033-295X.87.3.272

Dunn, B. R., Dunn, D. A., Languis, M., & Andrews, D. (1998). The relation of ERP components to complex memory processing. *Brain and cognition*, *36*(3), 355-376. doi:10.1006/brcg.1998.0998

Eimer, M., & Holmes, A. (2002). An ERP study on the time course of emotional face processing. *Neuroreport*, *13*(4), 427-431. doi: 10.1097/00001756-200203250-00013

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In Cole, J. (Ed.), *Nebraska symposium on motivation* (pp. 207-282). Lincoln, NB: University of Nebraska Press. Retrieve from https://www.paulekman.com/wp-content/uploads/2013/07/Universals-And-Cultural-Differences-In-Facial-Expressions-Of.pdf

Ekman, P. (1992). An argument for basic emotions. *Cognition & emotion*, *6*(3-4), 169-200. doi: 10.1080/02699939208411068

Ekman, P. (1994). Strong evidence for Universals in facial expressions: a reply to Russell's mistaken critique. *Psychological bulletin*, *115*(2), 268-287. doi: 10.1037//0033-2909.115.2.268

Ekman, P. (1999). Basic Emotions. In T. Dalgleish & M. J. Power (Eds.), *Handbook of Cognition and Emotion* (pp. 45-60). New York, NY: John Wiley & Sons Ltd.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, *17*(2), 124-129. doi: 10.1037/h0030377

Ekman,P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental psychology and nonverbal behavior*, *1*(1), 56-75. doi: 10.1007/BF01115465

Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System (FACS): A technique for the measurement of facial movements.* Palo Alto, CA: Consulting Psychologist Press.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, *1*(1), 49-98. Retrieve from http://boccignone.di.unimi.it/CompAff2011_files/EkmanFriesenSemiotica.pdf

Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science*, *164*(3875), 86-88. Retrieve from http://www.communicationcache.com/uploads/1/0/8/8/10887248/pan-cultural_elements_in_facial_displays_of_emotions.pdf

Ekman, P., Friesen, W. V., & Ellsworth, P. (1982). What is the relative contribution of facial behavior and contextual information to the judgment of emotion? In P.

Ekman (Ed.), *Emotion in the human face (*2nd ed*,* pp.111-127). New York: Cambridge University Press.

Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., ... & Ricci-Bitti, P. E. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of personality and social psychology*, *53*(4), 712-717.

Erthal, F. S., De Oliveira, L., Mocaiber, I., Pereira, M. G., Machado-Pinheiro, W., Volchan, E., & Pessoa, L. (2005). Load-dependent modulation of affective picture processing. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(4), 388-395. doi: 10.3758/CABN.5.4.388

Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition*, *44*(3), 227-240. doi:10.1016/0010-0277(92)90002-Y

Ethofer, T., Pourtois, G., & Wildgruber, D. (2006a). Investigating audiovisual integration of emotional signals in the human brain. *Progress in brain research*,*156*, 345-361. doi:10.1016/S0079-6123(06)56019-4

Ethofer, T., Anders, S., Erb, M., Droll, C., Royen, L., Saur, R., Reiterer, S., Grodd, W., & Wildgruber, D. (2006b). Impact of voice on emotional judgment of faces: An event-related fMRI study. *Human brain mapping*, *27*(9), 707-714. doi: 10.1002/hbm.20212

Fairhall, S. L., & Macaluso, E. (2009). Spatial attention can modulate audiovisual integration at multiple cortical and subcortical sites. *European Journal of Neuroscience*, *29*(6), 1247-1257. doi: 10.1111/j.1460-9568.2009.06688.x

Feyereisen, P., Malet, C., & Martin, Y. (1986). Is the faster processing of expressions of happiness modality-specific? In H. D. Ellis, M. A. Jeeves, F. Newcombe, & A. Young (Eds.), *Aspects of face processing* (pp. 349–355). Boston: Martinus Nijhoff.

Folstein, J. R., & van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology*, *45*(1), 152-170. doi: 10.1111/j.1469-8986.2007.00602.x

Foxe, J. J., Morocz, I. A., Murray, M. M., Higgins, B. A., Javitt, D. C., & Schroeder, C. E. (2000). Multisensory auditory–somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Cognitive Brain Research*, *10*(1), 77-83. doi:10.1016/S0926-6410(00)00024-0

Fugate, J. M. (2013). Categorical perception for emotional faces. *Emotion Review*, *5*(1), 84-89. doi: 10.1177/1754073912451350

Fujimura, T., Matsuda, Y. T., Katahira, K., Okada, M., & Okanoya, K. (2012). Categorical and dimensional perceptions in decoding emotional facial expressions. *Cognition & emotion*, *26*(4), 587-601. doi: 10.1080/02699931.2011.595391

Gehring, W. J., & Knight, R. T. (2000). Prefrontal–cingulate interactions in action monitoring. *Nature neuroscience*, *3*(5), 516-520. doi:10.1038/74899

Glowinski, D., Camurri, A., Volpe, G., Dael, N., & Scherer, K. (2008, June). Technique for automatic emotion recognition by body gesture analysis. In *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on* (pp. 1-6). IEEE.

Goldstone, R. L., & Hendrickson, A. T. (2010). Categorical perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(1), 69-78. doi: 10.1002/wcs.26

Grèzes, J., Pichon, S., & De Gelder, B. (2007). Perceiving fear in dynamic body expressions. *Neuroimage*, *35*(2), 959-967. doi:10.1016/j.neuroimage.2006.11.030

Gu, Y., Mai, X., & Luo, Y. J. (2013). Do bodily expressions compete with facial expressions? Time course of integration of emotional signals from the face and the body. *PloS one*, *8*(7), e66762. doi:10.1371/journal.pone.0066762.

Gunes, H., & Piccardi, M. (2006, August). A bimodal face and body gesture database for automatic analysis of human nonverbal affective behavior. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on* (Vol. 1, pp. 1148-1153). IEEE.

Gur, R. C., Schroeder, L., Turner, T., McGrath, C., Chan, R. M., Turetsky, B. I., ... & Gur, R. E. (2002). Brain activation during facial emotion processing. *Neuroimage*, *16*(3), 651-662. doi:10.1006/nimg.2002.1097

Hackley, S. A., Woldorff, M., & Hillyard, S. A. (1990). Cross-Modal Selective Attention Effects on Retinal, Myogenic, Brainstem, and Cerebral Evoked Potentials. *Psychophysiology*, *27*(2), 195-208. doi: 10.1111/j.1469-8986.1990.tb00370.x

Hadjikhani, N., Kveraga, K., Naik, P., & Ahlfors, S. P. (2009). Early (N170) activation of face-specific cortex by face-like objects. *Neuroreport*, *20*(4), 403. doi: 10.1097/WNR.0b013e328325a8e1

Hancock, P.A., Oron-Gilad, T., & Szalma, J.L. (2007). Elaborations of the multiple resource theory of attention. In: A.F. Kramer, D.A. Wiegmann, and A. Kirlik (Eds.). Attention: From theory to practice. (pp. 45-56). Oxford: Oxford University Press.

Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: an anger superiority effect. *Journal of personality and social psychology*, *54*(6), 917-924. doi: 10.1037/0022-3514.54.6.917

Harnad, S. (1987). Psychophysical and cognitive aspects of categorical perception: A critical overview. *Categorical perception: The groundwork of cognition*, (pp.1-52). New York: Cambridge university Press.

Hartikainen, K. M., Ogawa, K. H., & Knight, R. T. (2000). Transient interference of right hemispheric function due to automatic emotional processing. *Neuropsychologia*, *38*(12), 1576-1580. doi:10.1016/S0028-3932(00)00072-5

Harwood, N. K., Hall, L. J., & Shinkfield, A. J. (1999). Recognition of facial emotional expressions from moving and static displays by individuals with mental retardation. *American Journal on Mental Retardation*, *104*(3), 270-278. doi: 10.1352/0895-8017(1999)104<0270:ROFEEF>2.0.CO;2

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in cognitive sciences*, *4*(6), 223-233. doi:10.1016/S1364-6613(00)01482-0

Hein, G., & Knight, R. T. (2008). Superior temporal sulcus—it's my area: or is it?. *Cognitive Neuroscience, Journal of*, *20*(12), 2125-2136. doi: 10.1162/jocn.2008.20148

Hennenlotter, A. (2005). *Neural systems for recognising emotion from facial expressions* (Doctoral thesis, University of Regensburg, Munich, German). Retrieve from http://epub.uni-regensburg.de/10340/1/Hennenlotter_Dissertation_pdf.pdf)

Henson, R. N., Goshen-Gottstein, Y., Ganel, T., Otten, L. J., Quayle, A., & Rugg, M. D. (2003). Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cerebral cortex*, *13*(7), 793-805. *doi: 10.1093/cercor/13.7.793*

Herrmann, M. J., Ehlis, A. C., Ellgring, H., & Fallgatter, A. J. (2005). Early stages (P100) of face perception in humans as measured with event-related potentials (ERPs). *Journal of neural transmission*, *112*(8), 1073-1081. doi: 10.1007/s00702-004-0250-8

Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behavior*, *21*(4), 241-257. doi: 10.1023/A:1024952730333

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*(108), 177-180. doi: 10.1126/ science.182.4108.177

Hoffmann, H., Kessler, H., Eppel, T., Rukavina, S., & Traue, H. C. (2010). Expression intensity, gender and facial emotion recognition: Women recognize only subtle facial emotions better than men. *Acta psychologica*,*135*(3), 278-283. doi:10.1016/j.actpsy.2010.07.012

Holmes, A., Vuilleumier, P., & Eimer, M. (2003). The processing of emotional facial expression is gated by spatial attention: evidence from event-related brain potentials. *Cognitive Brain Research*, *16*(2), 174-184. doi:10.1016/S0926-6410(02)00268-9

Hopfinger, J. B., Buonocore, M. H., & Mangun, G. R. (2000). The neural mechanisms of top-down attentional control. *Nature neuroscience*, *3*(3), 284-291. doi:10.1038/72999

Horstmann, G., & Bauland, A. (2006). Search asymmetries with real faces: testing the anger-superiority effect. *Emotion*, *6*(2), 193-207. doi: 10.1037/1528-3542.6.2.193

Ishai, A., Pessoa, L., Bikle, P. C., & Ungerleider, L. G. (2004). Repetition suppression of faces is modulated by emotion. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(26), 9827-9832. doi: 10.1073/pnas.0403559101

Itier, R. J., & Taylor, M. J. (2004). N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cerebral cortex*, *14*(2), 132-142. *doi: 10.1093/cercor/bhg111*

Izard, C. E. (1971). *The face of emotion.* East Norwalk, CT, US: Appleton-Century-Crofts.

Jack, R. E., Garrod, O. G., & Schyns, P. G. (2014). Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Current biology*, *24*(2), 187-192. doi:10.1016/j.cub.2013.11.064

Kahneman, D. (1973). *Attention and Effort.* Englewood Cliffs, NJ: Prentice-Hall

Keidel, W. D., & Spreng, M. (1965). Audiometric aspects and multisensory power-functions of electronically averaged slow evoked cortical responses in man. *Acta Oto-Laryngologica*, *59*(2-6), 201-210. doi: 10.3109/00016486509124553

Kim, S. Y., Kim, M. S., & Chun, M. M. (2005). Concurrent working memory load can reduce distraction. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(45), 16524-16529. doi: 10.1073/pnas.0505454102

Kirita, T., & Endo, M. (1995). Happy face advantage in recognizing facial expressions. *Acta psychologica*, *89*(2), 149-163. doi:10.1016/0001-6918(94)00021-8

Kirouac, G., & Doré, F. Y. (1983). Accuracy and latency of judgment of facial expressions of emotions. *Perceptual and motor skills*, *57*(3), 683-686. doi: 10.2466/pms.1983.57.3.683

Kleinsmith, A., & Bianchi-Berthouze, N. (2013). Affective body expression perception and recognition: A survey. *Affective Computing, IEEE Transactions on*, *4*(1), 15-33. doi: 10.1109/T-AFFC.2012.16

Kotsoni, E., de Haan, M., & Johnson, M. H. (2001). Categorical perception of facial expressions by 7-month-old infants. *Perception*, *30*(9), 1115-1126. doi: 10.1068/p3155

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological psychology*, *84*(3), 394-421. doi:10.1016/j.biopsycho.2010.03.010

Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., & Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. *Neuroimage*, *37*(4), 1445-1456. doi:10.1016/j.neuroimage.2007.06.020

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

Lane, R. D., Chua, P. M., & Dolan, R. J. (1999). Common effects of emotional valence, arousal and attention on neural activation during visual processing of pictures. *Neuropsychologia*, *37*(9), 989-997. doi:10.1016/S0028-3932(99)00017-2

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: brain mechanisms and psychophysiology. *Biological psychiatry*,*44*(12), 1248-1263. doi:10.1016/S0006-3223(98)00275-3

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 451-468. doi: 10.1037//0096-1523.21.3.451

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in cognitive sciences*, *9*(2), 75-82. doi:10.1016/j.tics.2004.12.004

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, *56*(2), 183-197. doi: 10.3758/BF03213897

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339-354. doi: 10.1037/0096-3445.133.3.339

Lee, T. H., Choi, J.S., & Cho, Y.S. (2012). Contex modulation of facial emotion perception differed by individual difference. *PloS ONE, 7*(3), e32987. doi: 10.1371/journal.pone.0032987.

Lee, S. Y., Kang, J. I., Lee, E., Namkoong, K., & An, S. K. (2011). Differential priming effect for subliminal fear and disgust facial expressions. *Attention, Perception, & Psychophysics*, *73*(2), 473-481. doi: 10.3758/s13414-010-0032-3

Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: a meta-analysis of experimental emotion elicitations. *Psychological bulletin*, *137*(5), 834-855. doi: 10.1037/a0024244

Lim, S. L., & Pessoa, L. (2008). Affective learning increases sensitivity to graded emotional faces. *Emotion*, *8*(1), 96-103.doi: 10.1037/1528-3542.8.1.96

Liu, K., & Picard, R. W. (2003, April). *Subtle expressivity in a robotic computer*. Paper presented at CHI 2003 Workshop on Subtle Expressiveness in Characters and Robots. Fort Lauderdale,Florida, US.

Liu, T., Pinheiro, A. P., Deng, G., Nestor, P. G., McCarley, R. W., & Niznikiewicz, M. A. (2012). Electrophysiological insights into processing nonverbal emotional vocalizations. *Neuroreport*, *23*(2), 108-112. doi: 10.1097/WNR.0b013e32834ea757

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique.* Cambridge, MA: MIT Press.

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*(3), 291-308. doi: 10.1111/j.1469-8986.1994.tb02218.x

Luck, S. J., & Hillyard, S. A. (2000). The operation of selective attention at multiple stages of processing: Evidence from human and monkey electrophysiology. In M.S. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 687-710). Cambridge, MA: MIT Press.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in cognitive sciences*, *4*(11), 432-440. doi:10.1016/S1364-6613(00)01545-X

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces-KDEF*. CD-ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, Stockholm, Sweden*. ISBN 91-630-7164-9.

Marsh, A. A., Ambady, N., & Kleck, R. E. (2005). The effects of fear and anger facial expressions on approach-and avoidance-related behaviors. *Emotion*, *5*(1), 119-124. doi: 0.1037/1528-3542.5.1.119

Martinez, A., & Du, S. (2012). A model of the perception of facial expressions of emotion by humans: Research overview and perspectives. *The Journal of Machine Learning Research*, *13*(1), 1589-1608. Retrieve from http://www.jmlr.org/papers/volume13/martinez12a/martinez12a.pdf

McAndrew, F. T. (1986). A cross-cultural study of recognition thresholds for facial expressions of emotion. *Journal of Cross-Cultural Psychology*, *17*(2), 211-224. doi: 10.1177/0022002186017002005

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746-748. doi: 10.1038/264746a0

Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(45), 16518-16523. doi: 10.1073/pnas.0507650102

Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., & Kramer, A. F. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Cognitive Brain Research*, *12*(3), 467-473. doi:10.1016/S0926-6410(01)00076-3

Mitchell, J. F., Sundberg, K. A., & Reynolds, J. H. (2007). Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron*, *55*(1), 131-141. doi:10.1016/j.neuron.2007.06.018

Miyoshi, M., Katayama, J. I., & Morotomi, T. (2004). Face-specific N170 component is modulated by facial expressional change. *Neuroreport*, *15*(5), 911-914. doi: 10.1097/00001756-200404090-00035

Mobbs, D., Weiskopf, N., Lau, H. C., Featherstone, E., Dolan, R. J., & Frith, C. D. (2006). The Kuleshov Effect: the influence of contextual framing on emotional attributions. *Social cognitive and affective neuroscience*, *1*(2), 95-106. *doi: 10.1093/scan/nsl014*

Mondloch, C. J. (2012). Sad or fearful? The influence of body posture on adults' and children's perception of facial displays of emotion. *Journal of experimental child psychology*, *111*(2), 180-196. doi:10.1016/j.jecp.2011.08.003

Mondloch, C. J., Horner, M., & Mian, J. (2013a). Wide eyes and drooping arms: Adult-like congruency effects emerge early in the development of sensitivity to emotional faces and body postures. *Journal of experimental child psychology*, *114*(2), 203-216. doi:10.1016/j.jecp.2012.06.003

Mondloch, C. J., Nelson, N. L., & Horner, M. (2013b). Asymmetries of influence: differential effects of body postures on perceptions of emotional facial expressions. *PloS one*, *8*(9), e73605. doi:10.1371/journal.pone.0073605

Montagne, B., Kessels, R. P., De Haan, E. H., & Perrett, D. I. (2007). The emotion recognition rask: a paradigm to measure the perception of facial emotional expressions at different intensities. *Perceptual and motor skills*, *104*(2), 589-598. doi: 10.2466/pms.104.2.589-598

Morris, J. S., Friston, K. J., Buchel, C., Frith, C. D., Young, A. W., Calder, A. J., & Dolan, R. J. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain*, *121*(1), 47-57. doi: 10.1093/brain/121.1.47

Mourao-Miranda, J., Volchan, E., Moll, J., de Oliveira-Souza, R., Oliveira, L., Bramati, I., ... & Pessoa, L. (2003). Contributions of stimulus valence and arousal to visual activation during emotional perception. *Neuroimage*, *20*(4), 1955-1963. doi:10.1016/j.neuroimage.2003.08.011

Mozolic, J. L., Hugenschmidt, C. E., Peiffer, A. M., & Laurienti, P. J. (2008). Modality-specific selective attention attenuates multisensory integration. *Experimental brain research*, *184*(1), 39-52. doi: 10.1007/s00221-007-1080-3

Müller, V. I., Habel, U., Derntl, B., Schneider, F., Zilles, K., Turetsky, B. I., & Eickhoff, S. B. (2011). Incongruence effects in crossmodal emotional integration. *Neuroimage*, *54*(3), 2257-2266. doi:10.1016/j.neuroimage.2010.10.047

Nash, A. J., & Williams, C. S. (1982). Effects of preparatory set and task demands on auditory event-related potentials. *Biological psychology*, *15*(1), 15-31. doi:10.1016/0301-0511(82)90028-X

Nguyen, V. T., & Cunnington, R. (2014). The superior temporal sulcus and the N170 during face processing: Single trial analysis of concurrent EEG–fMRI. *Neuroimage*, *86*, 492-502. doi:10.1016/j.neuroimage.2013.10.047

Niit, T., & Valsiner, J. (1977). Recognition of facial expressions: An experimental investigation of Ekman's model. *Tartu Riikliku Ülikooli Toimetised: Trudy po Psikhologii*, *429,* 85-107.

Oatley, K., & Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition and emotion*, *1*(1), 29-50. doi: 10.1080/02699938708408362

Oh, S. H., & Kim, M. S. (2004). The role of spatial working memory in visual search efficiency. *Psychonomic Bulletin & Review*, *11*(2), 275-281. doi: 10.3758/BF03196570

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological review*, *108*(3), 483 doi: 10.1037//0033-295x.108.3.483.

Okon-Singer, H., Tzelgov, J., & Henik, A. (2007). Distinguishing between automaticity and attention in the processing of emotionally significant stimuli. *Emotion*, *7*(1), 147-157. doi: 10.1037/1528-3542.7.1.147

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in cognitive sciences*, *11*(12), 520-527. doi:10.1016/j.tics.2007.09.009

Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: an integrative review of ERP findings. *Biological psychology*, *77*(3), 247-265. doi:10.1016/j.biopsycho.2007.11.006

Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions?. *Psychological review*, *97*(3), 315. Retrieve from http://www.communicationcache.com/uploads/1/0/8/8/10887248/whats_basic_about_basic_emotions.pdf

Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia*, *45*(1), 75-92. doi:10.1016/j.neuropsychologia.2006.04.025

Panayiotou, G. (2008). Emotional dimensions reflected in ratings of affective scripts. *Personality and Individual Differences*, *44*(8), 1795-1806. doi:10.1016/j.paid.2008.02.006

Paus, T. S. (2001). Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Nature Reviews Neuroscience*, *2*(6), 417-424. doi: 10.1038/35077500

Park, S., Kim, M. S., & Chun, M. M. (2007). Concurrent working memory load can facilitate selective attention: evidence for specialized load. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(5), 1062-1075. doi: 10.1037/0096-1523.33.5.1062

Pereira, M. G., Volchan, E., de Souza, G. G. L., Oliveira, L., Campagnoli, R. R., Pinheiro, W. M., & Pessoa, L. (2006). Sustained and transient modulation of performance induced by emotional picture viewing. *Emotion*, *6*(4), 622-634. doi: 10.1037/1528-3542.6.4.622

Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *British Medical Journal, 316*(7139), 1236-1238. Retrieve from http://www.stat.wvu.edu/~gmerovic/QuantEcol_files/WMAN633_Perneger1998_Bonferroni_adjustments.pdf

Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: from a'low road'to'many roads' of evaluating biological significance. *Nature Reviews Neuroscience*, *11*(11), 773-783. doi:10.1038/nrn2920

Pessoa, L., & Adolphs, R. (2011). Emotion and the brain: multiple roads are better than one. *Nature Reviews Neuroscience*, *12*(7), 425-425. doi:10.1038/nrn2920-c2

Pessoa, L., McKenna, M., Gutierrez, E., & Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences*, *99*(17), 11458-11463. doi: 10.1073/pnas.172403899

Pessoa, L., Padmala, S., & Morland, T. (2005). Fate of unattended fearful faces in the amygdala is determined by both attentional resources and cognitive modulation. *Neuroimage*, *28*(1), 249-255. doi:10.1016/j.neuroimage.2005.05.048

Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., ... & Gray, J. A. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature*, *389*(6650), 495-498. doi:10.1038/39051

Phillips, M. L., Williams, L. M., Heining, M., Herba, C. M., Russell, T., Andrew, C., ... & Young, A. W. (2004). Differential neural responses to overt and covert presentations of facial expressions of fear and disgust. *Neuroimage*, *21*(4), 1484-1496. doi:10.1016/j.neuroimage.2003.12.013

Piech, R. M., McHugo, M., Smith, S. D., Dukic, M. S., Van Der Meer, J., Abou-Khalil, B., & Zald, D. H. (2010). Fear-enhanced visual search persists after amygdala lesions. *Neuropsychologia*, *48*(12), 3430-3435. doi:10.1016/j.neuropsychologia.2010.07.009

Piwek, L., Pollick, F., & Petrini, K. (2015). Audiovisual integration of emotional signals from others' social interactions. *Frontiers in psychology*, *9*, 116-125. doi: 10.3389/fpsyg.2015.00611

Plutchik, R. (1962). *The Emotions: Facts, Theories, and a New Model,* New York: Random House.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In Plutchik, R., Kellerman, H. (Eds.), *Emotion, theory, research, and experience* (pp. 3–33). New York: Acacemic Press.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology*, *118*(10), 2128-2148. doi:10.1016/j.clinph.2007.04.019

Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., & Crommelinck, M. (2000). The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport*, *11*(6), 1329-1333. Retrieve from http://spitswww.uvt.nl/fsw/psychologie/cognitive-neuropsychology/pubs/jean/27.pdf

Redcay, E. (2008). The superior temporal sulcus performs a common function for social and speech perception: implications for the emergence of autism. *Neuroscience & Biobehavioral Reviews*, *32*(1), 123-142. doi:10.1016/j.neubiorev.2007.06.004

Righart, R., & de Gelder, B. (2008a). Rapid influence of emotional scenes on encoding of facial expressions: an ERP study. *Social Cognitive and Affective Neuroscience*, *3*(3), 270-278. *doi: 10.1093/scan/nsn021*

Righart, R., & De Gelder, B. (2008b). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(3), 264-272. doi: 10.3758/CABN.8.3.264

Roether, C. L., Omlor, L., & Giese, M. A. (2010). Features in the recognition of emotions from dynamic bodily expression. In G. S. Masson & U. J. Ilg (Eds). *Dynamics of Visual Motion Processing* (pp. 313-340). New York: Springer US. doi: 10.1007/978-1-4419-0781-3

Rosenberg, E. L., & Ekman, P. (1994). Coherence between expressive and experiential systems in emotion. *Cognition & Emotion*, *8*(3), 201-229. doi: 10.1080/02699939408408938

Rothman, K. J. (1990). No adjustments are needed for multiple comparisons. *Epidemiology, 1*(1), 43-46. Retrieve from http://www.jstor.org.ezproxy.massey.ac.nz/stable/pdf/20065622.pdf?_=1468279771197

Rozin, P., Lowery, L., & Ebert, R. (1994). Varieties of disgust faces and the structure of disgust. *Journal of personality and social psychology*, *66*(5), 870-881. doi: 10.1037//0022-3514.66.5.870

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, *39*(6), 1161-1178. Retrieve from http://web.b.ebscohost.com.ezproxy.massey.ac.nz/ehost/pdfviewer/pdfviewer?vid=3&sid=8de7e097-9e75-4a08-8645-fdfd8fde5874%40sessionmgr120&hid=115

Russell, J. A. (1994). Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies. *Psychological bulletin*,*115*(1), 102-141. Retrieve from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.85.4287&rep=rep1&type=pdf

Russell, J. A. (1997). Reading emotion from and into faces: resurrecting a dimensional-contextual perspective. In Russell J.A., Fernandez-Dols J. M. (Eds.), *The psychology of facial expression* (pp. 295-320). Cambridge, UK: Cambridge Univeristy Press.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological review*, *110*(1), 145-172. doi: 10.1037/0033-295X.110.1.145

Russell, J. A., & Bullock, M. (1985). Multidimensional scaling of emotional facial expressions: similarity from preschoolers to adults. *Journal of Personality and Social Psychology*, *48*(5), 1290-1298. doi: 10.1037/0022-3514.48.5.1290

Russell, J. A., Bachorowski, J. A., & Fernández-Dols, J. M. (2003). Facial and vocal expressions of emotion. *Annual review of psychology*, *54*(1), 329-349. doi: 10.1146/annurev.psych.54.101601.145102

Sander, D., Grandjean, D., Kaiser, S., Wehrle, T., & Scherer, K. R. (2007). Interaction effects of perceived gaze direction and dynamic facial expression: Evidence for appraisal theories of emotion. *European Journal of Cognitive Psychology*, *19*(3), 470-480. doi: 10.1080/09541440600757426

Saville, D. J. (1990). Multiple comparison procedures: the practical solution. *The American Statistician,* *44*(2), 174-180. Retrieve from http://www.jstor.org.ezproxy.massey.ac.nz/stable/pdf/2684163.pdf

Scherer, K. R. (1981). Speech and emotional states. In Darby, J. K. (Ed.), *Speech evaluation in psychiatry* (pp. 189-220). New York: Grune and Stratton.

Scherer, K. R. (2000). Psychological models of emotion. In J. C. Borod (Ed), *The neuropsychology of emotion* (pp.137-162). New York, NY, US: Oxford University Press.

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech communication*, *40*(1), 227-256. doi:10.1016/S0167-6393(02)00084-5

Scherer, K. R., Ladd, D. R., & Silverman, K. E. (1984). Vocal cues to speaker affect: Testing two models. *The Journal of the Acoustical Society of America*, *76*(5), 1346-1356. doi: 10.1121/1.391450

Schindler, K., Van Gool, L., & de Gelder, B. (2008). Recognizing emotions expressed by body pose: A biologically inspired neural model. *Neural networks*,*21*(9), 1238-1246. doi:10.1016/j.neunet.2008.05.003

Schlosberg, H. (1954). Three dimensions of emotion. *Psychological review*,*61*(2), 81-88. doi: 10.1037/h0054570

Schröder, M. (2003). Experimental study of affect bursts. *Speech communication*, *40*(1), 99-116. doi:10.1016/S0167-6393(02)00078-X

Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2003). Attention and emotion: an ERP analysis of facilitated emotional stimulus processing. *Neuroreport*, *14*(8), 1107-1110. doi: 10.1097/00001756-200306110-00002

Schupp, H. T., Stockburger, J., Codispoti, M., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2007). Selective visual attention to emotion. *The Journal of Neuroscience*, *27*(5), 1082-1089. Retrieve from http://www.mnf.uni-greifswald.de/fileadmin/psychologie/lehrstuhl_klin/artikel/SchuppEtAl_JNeuroSci2007.pdf

Sedgwick, P. (2014). Multiple hypothesis testing and Bonferroni's correction. *British Medical Journal. 349*:g6284. doi:10.1136/bmj.g6284

Shan, C., Gong, S., & McOwan, P. W. (2007, September). Beyond Facial Expressions: Learning Human Emotion from Body Gestures. Paper presented at the British Machine Vision Conference, Warwick, UK.

Shariff, A. F., & Tracy, J. L. (2011). What are emotion expressions for?. *Current Directions in Psychological Science*, *20*(6), 395-399. doi: 10.1177/0963721411424739

Shibata, T., Nishijo, H., Tamura, R., Miyamoto, K., Eifuku, S., Endo, S., & Ono, T. (2002). Generators of visual evoked potentials for faces and eyes in the human brain as determined by dipole localization. *Brain topography*,*15*(1), 51-63. doi: 10.1023/A:1019944607316

Silver, M. A., Ress, D., & Heeger, D. J. (2007). Neural correlates of sustained spatial attention in human early visual cortex. *Journal of neurophysiology*, *97*(1), 229-237. doi: 10.1152/jn.00677.2006

Simpson, J. R., Öngür, D., Akbudak, E., Conturo, T. E., Ollinger, J. M., Snyder, A. Z., ... & Raichle, M. E. (2006). The emotional modulation of cognitive processing: an fMRI study. doi: 10.1162/089892900564019

Spence, C., Senkowski, D., & Röder, B. (2009). Crossmodal processing. *Experimental brain research*, *198*(2-3), 107-111. doi: 10.1007/s00221-009-1973-4

Spreckelmeyer, K. N., Kutas, M., Urbach, T. P., Altenmüller, E., & Münte, T. F. (2006). Combined perception of emotion in pictures and musical sounds. *Brain research*, *1070*(1), 160-170. doi:10.1016/j.brainres.2005.11.075

Spreng, M. (1980). Influence of impulsive and fluctuating noise upon physiological excitations and short-time readaptation. *Scandinavian audiology. Supplementum*, (Supplement 12), 299-306.

Sprengelmeyer, R., Rausch, M., Eysel, U. T., & Przuntek, H. (1998). Neural structures associated with recognition of facial expressions of basic emotions. *Proceedings of the Royal Society of London B: Biological Sciences*, *265*(1409), 1927-1931. doi: 10.1098/rspb.1998.0522

Stekelenburg, J. J., & de Gelder, B. (2004). The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. *Neuroreport*, *15*(5), 777-780. Retrieve from http://www.ece.uvic.ca/~bctill/papers/facerec/Stekelenburg2004a.pdf

Sur, S., & Sinha, V. K. (2009). Event-related potential: An overview. *Industrial psychiatry journal*, *18*(1), 70-73. doi: 10.4103/0972-6748.57865

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*(1), 152-162. doi:10.1016/j.neuropsychologia.2006.05.001

Suzuki, N., & Bartneck, C. (2003, April). Subtle expressivity for characters and robots. Paper presented at *CHI'03 Extended Abstracts on Human Factors in Computing Systems*, Fort Lauderdale,Florida, US . doi: 10.1145/765891.766150

Suzuki, A., Hoshino, T., Shigemasu, K., & Kawamura, M. (2006). Disgust-specific impairment of facial expression recognition in Parkinson's disease. *Brain*, *129*(3), 707-717. doi: *10.1093/brain/awl011*

Takehara, T., & Suzuki, N. (1997). Morphed images of basic emotional expressions: Ratings on Russell's bipolar field. *Perceptual and motor skills*, *85*(3), 1003-1010. doi: 10.2466/pms.1997.85.3.1003

Talsma, D., & Woldorff, M. G. (2005). Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. *Cognitive Neuroscience, Journal of*, *17*(7), 1098-1114. doi: 10.1162/0898929054475172

Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration?. *Cerebral Cortex*, *17*(3), 679-690. *doi: 10.1093/cercor/bhk016*

Tanaka-Matsumi, J., Attivissimo, D., Nelson, S., & D'Urso, T. (1995). Context effects on the judgment of basic emotions in the face. *Motivation and Emotion*, *19*(2), 139-155. doi: 10.1007/BF02250567

Thierry, G., Martin, C. D., Downing, P. E., Pegna, A. J. (2007). Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nature Neuroscience, 10*(4), 505-511. doi: 10.1018/nn1864

Thoma, P., Bauser, D. S., & Suchan, B. (2013). BESST (Bochum Emotional Stimulus Set)—A pilot validation study of a stimulus set containing emotional bodies and faces from frontal and averted views. *Psychiatry research*,*209*(1), 98-109. doi:10.1016/j.psychres.2012.11.012

Tipples, J., & Sharma, D. (2000). Orienting to exogenous cues and attentional bias to affective pictures reflect separate processes. *British Journal of Psychology*, *91*(1), 87-97. doi: 10.1348/000712600161691

Tomkins, S. S. (1962). *Affect, imagery, consciousness: Vol. I: The positive affects.* New York: Springer.

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In N. J. Castellan & F. Restle (Eds.), *Cognitive theory*, (vol. *3*, pp.200-239). Hiilsdale, NJ: Erlbaum.

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, *12*(4), 242-248.

Treisman, A. (1964). Monitoring and storage of irrelevant messages in selective attention. *Journal of Verbal Learning and Verbal Behavior*, *3*(6), 449-459. doi:10.1016/S0022-5371(64)80015-3

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, *12*(1), 97-136. doi:10.1016/0010-0285(80)90005-5

Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(3), 459-478. Retrieve from https://www.semanticscholar.org/paper/Conjunction-Search-Revisited-Treisman-Sato/bf1ec479aabf41827a30fece1e66abb0cbd2fe57/pdf

Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review*, *93*(3), 239-257. doi: 10.1037/0033-295X.93.3.239

Tsao, D. Y., Freiwald, W. A., Tootell, R. B., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science*, *311*(5761), 670-674. doi: 10.1126/science.1119983

Tsuchiya, N., Moradi, F., Felsen, C., Yamazaki, M., & Adolphs, R. (2009). Intact rapid detection of fearful faces in the absence of the amygdala. *Nature neuroscience*, *12*(10), 1224-1225. doi: 10.1038/nn.2380

van de Riet, W. A., & de Gelder, B. (2008). Watch the face and look at the body!. *Netherlands Journal of Psychology*, *64*(4), 143-151. doi: 10.1007/BF03076417

van den Stock, J., Righart, R., & De Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, *7*(3), 487-494.doi: 10.1037/1528-3542.7.3.487

Volkova, E. P., Mohler, B. J., Dodds, T. J., Tesch, J., & Bülthoff, H. H. (2014). Emotion categorization of body expressions in narrative scenarios. *Frontiers in Psychology*, *5*, 623. doi: 10.3389/fpsyg.2014.00623

Vroomen, J., Bertelson, P., & de Gelder, B. (2001a). Directing spatial attention towards the illusory location of a ventriloquized sound. *Acta Psychologica*,*108*(1), 21-33. doi:10.1016/S0001-6918(00)00068-8

Vroomen, J., Bertelson, P., & de Gelder, B. (2001b). The ventriloquist effect does not depend on the direction of automatic visual attention. *Perception & psychophysics*, *63*(4), 651-659. doi: 10.3758/BF03194427

Vroomen, J., Driver, J., & De Gelder, B. (2001c). Is cross-modal integration of emotional expressions independent of attentional resources?. *Cognitive, Affective, & Behavioral Neuroscience*, *1*(4), 382-387. doi: 10.3758/CABN.1.4.382

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron*, *30*(3), 829-841. doi:10.1016/S0896-6273(01)00328-2

Wang, L., Fu, S., Feng, C., Luo, W., Zhu, X., & Luo, Y. J. (2012). The neural processing of fearful faces without attention and consciousness: An event-related potential study. *Neuroscience letters*, *506*(2), 317-321. doi:10.1016/j.neulet.2011.11.034

Wallbott, H. G. (1998). Bodily expression of emotion. *European journal of social psychology*, *28*(6), 879-896. doi: 10.1002/(SICI)1099-0992(1998110)28:6<879::AID-EJSP901>3.0.CO;2-W

Watanabe, S., Kakigi, R., & Puce, A. (2003). The spatiotemporal dynamics of the face inversion effect: a magneto-and electro-encephalographic study. *Neuroscience*, *116*(3), 879-895. doi:10.1016/S0306-4522(02)00752-2

Weissman, D. H., Gopalakrishnan, A., Hazlett, C. J., & Woldorff, M. G. (2005). Dorsal anterior cingulate cortex resolves conflict from distracting stimuli by boosting attention toward relevant events. *Cerebral Cortex*, *15*(2), 229-237. *doi: 10.1093/cercor/bhh125*

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 449-455. doi: 10.1518/001872008X288394

Willadsen-Jensen, E. C., & Ito, T. A. (2006). Ambiguity and the timecourse of racial perception. *Social Cognition*, *24*(5), 580-606. doi: 10.1521/soco.2006.24.5.580

Yik, M., Widen, S. C., & Russell, J. A. (2013). The within-subjects design in the study of facial expressions. *Cognition & emotion*, *27*(6), 1062-1072. doi: 10.1080/02699931.2013.763769

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (1997). Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. *Cognition*, *63*(3), 271-313. doi:10.1016/S0010-0277(97)00003-6

Appendix A ANOVA Results for the Current Experiments

Table 1. ANOVA Results for Accuracy in Experiment 1

Source	df	Mean Square	F	Sig	Partial Eta Squared
Main effects					
FA	3.11	81813.24	82.43	0.000	0.760
SIM	3	46563.38	296.60	0.000	0.919
Interaction effects					
*FA*SIM*	7.87	7729.17	30.16	0.000	0.537

Note: FA: Facial expression (anger, disgust, fear, happiness, sadness, surprise)

SIM: Perceptual similarity (full similarity, high similarity, low similarity, isolated facial expression)

Table 2. ANOVA Results for Accuracy in Experiment 2

Source	df	Mean Square	F	Sig	Partial Eta Squared
Main effects					
FA	3.54	31402.35	25.48	0.000	0.526
SIM	1.14	76255.74	28.67	0.000	0.555
Interaction effects					
*FA*SIM*	5.36	5352.84	11.04	0.000	0.324

Note: FA: Facial expression (anger, disgust, fear, happiness, sadness, surprise)

SIM: Perceptual similarity (full similarity, high similarity, low similarity, isolated facial expression)

Table 3. ANOVA Results for Accuracy in Experiment 3

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
FA	1.00	9082.18	4.31	0.046	0.126
Within-subjects effects					
CON	1.39	76721.43	45.75	0.000	0.604
INT	1.17	16446.56	15.74	0.000	0.344
Interaction effects					
*FA*CON*INT*	3.13	640.68	4.24	0.007	0.124
*FA*CON*	1.39	10791.11	6.43	0.008	0.177
*FA*INT*	1.17	7683.45	7.35	0.008	0.197
*CON*INT*	3.13	1653.12	10.93	0.000	0.267

Note: FA: Facial expression (disgust, fear)

CON: Consistency condition (congruent condition, incongruent condition, isolated facial expression)

INT: Level of intensity (40%, 60%, 80%, 100%)

Table 4. ANOVA Results for Consistency Effects in Experiment 3

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
EMO	1	26593.95	7.01	0.013	0.189
Within-subjects effects					
INT	2.24	750.17	6.20	0.002	0.171
Interaction effects					
*FA*INT*	2.24	1299.56	10.74	0.00	0.26

Note: EMO: Emotion (disgust, fear)

INT: Level of intensity (40%, 60%, 80%, 100%)

Table 5. ANOVA Results for Facilitation Effects in Experiment 3

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
EMO	1	970.75	0.48	0.494	0.016
Within-subjects effects					
INT	1.80	4953.91	15.64	0.000	0.343
Interaction effects					
*FA*INT*	1.80	1653.97	5.22	0.011	0.148

Note: EMO: Emotion (disgust, fear)

INT: Level of intensity (40%, 60%, 80%, 100%)

Table 6. ANOVA Results for Interference Effects in Experiment 3

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
EMO	1	17402.78	14.89	0.001	0.332
Within-subjects effects					
INT	2.01	2450.22	8.51	0.001	0.221
Interaction effects					
*FA*INT*	2.01	66.60	0.23	0.796	0.008

Note: EMO: Emotion (disgust, fear)

INT: Level of intensity (40%, 60%, 80%, 100%)

Table 7. ANOVA Results for Accuracy in Experiment 4

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
FA	1	591.59	0.49	0.491	0.023
Within-subjects effects					
CON	1.18	60181.61	24.16	0.000	0.535
INT	1.23	14369.46	9.96	0.002	0.322
Interaction effects					
*FA*CON*INT*	1.74	216.77	0.54	0.563	0.025
*FA*CON*	1.18	2661.21	1.07	0.324	0.048
*FA*INT*	1.23	12189.02	8.45	0.005	0.287
*CON*INT*	1.74	2831.85	7.04	0.004	0.251

Note: FA: Facial expression (disgust, fear)

CON: Consistency condition (congruent condition, incongruent condition, isolated facial expression)

INT: Level of intensity (40%, 60%, 80%, 100%)

Table 8. ANOVA Results for Consistency Effects in Experiment 4

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
EMO	1	3685.45	0.69	0.414	0.032
Within-subjects effects					
INT	3	77.47	0.95	0.421	0.043
Interaction effects					
*FA*INT*	3	118.42	1.45	0.24	0.07

Note: EMO: Emotion (disgust, fear)

INT: Level of intensity (40%, 60%, 80%, 100%)

Table 9. ANOVA Results for Facilitation Effects in Experiment 4

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
EMO	1	5533.52	4.51	0.046	0.177
Within-subjects effects					
INT	1.34	4875.52	7.21	0.007	0.256
Interaction effects					
*FA*INT*	1.34	155.77	0.23	0.705	0.011

Note: EMO: Emotion (disgust, fear)

INT: Level of intensity (40%, 60%, 80%, 100%)

Table 10. ANOVA Results for Interference Effects in Experiment 4

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
EMO	1	187.13	0.08	0.777	0.004
Within-subjects effects					
INT	1.26	6359.58	8.45	0.005	0.287
Interaction effects					
*FA*INT*	1.26	450.31	0.60	0.483	0.028

Note: EMO: Emotion (disgust, fear)

INT: Level of intensity (40%, 60%, 80%, 100%)

Table 11. ANOVA Results for d' in Experiment 5

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
FA	1	8.76	3.63	0.072	0.160
Within-subjects effects					
CON	2	0.87	2.44	0.109	0.114
LOAD	1	22.99	31.35	0.000	0.623
Interaction effects					
*FA*CON*LOAD*	2	0.08	0.37	0.694	0.019
*FA*CON*	2	0.60	1.92	0.161	0.092
*FA*LOAD*	1	1.13	1.34	0.230	0.075
*CON*LOAD*	2	0.30	1.34	0.273	0.066

Note: FA: Facial expression (disgust, fear)

CON: Consistency condition (congruent condition, incongruent condition, isolated facial expression)

LOAD: Perceptual load (low load, high load)

Table 12. ANOVA Results for Accuracy in Experiment 5

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
FA	1	0.72	0.00	0.976	0.000
Within-subjects effects					
CON	1.30	1.26	17.86	0.000	0.484
LOAD	1	0.01	0.00	0.996	0.000
Interaction effects					
*FA*CON*LOAD*	2	1010.79	8.20	0.001	0.301
*FA*CON*	1.30	0.69	9.88	0.002	0.342
*FA*LOAD*	1	446.52	1.23	0.280	0.061
*CON*LOAD*	2	726.33	5.89	0.006	0.237

Note: FA: Facial expression (disgust, fear)

CON: Consistency condition (congruent condition, incongruent condition, isolated facial expression)

LOAD: Perceptual load (low load, high load)

Table 13. ANOVA Results for Reaction Times in Experiment 5

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
FA	1	147121.11	0.68	0.422	0.034
Within-subjects effects					
CON	2	44255.80	4.93	0.013	0.206
LOAD	1	615211.30	10.26	0.005	0.351
Interaction effects					
*FA*CON*LOAD*	1.39	1564.28	0.09	0.840	0.005
*FA*CON*	2	15237.34	1.70	0.197	0.082
*FA*LOAD*	1	232715.73	3.88	0.064	0.170
*CON*LOAD*	1.39	1340.49	0.08	0.856	0.004

Note: FA: Facial expression (disgust, fear)

CON: Consistency condition (congruent condition, incongruent condition, isolated facial expression)

LOAD: Perceptual load (low load, high load)

Table 14. ANOVA Results for d' in Experiment 6A

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
FA	1	0.35	0.29	0.598	0.014
Within-subjects effects					
CON	2	0.38	1.41	0.257	0.066
LOAD	1	39.30	65.19	0.000	0.765
Interaction effects					
*FA*CON*LOAD*	2	0.07	0.28	0.757	0.014
*FA*CON*	2	0.12	0.45	0.640	0.022
*FA*LOAD*	1	0.50	0.83	0.373	0.040
*CON*LOAD*	2	0.13	0.51	0.605	0.025

Note: FA: Facial expression (disgust, fear)

CON: Consistency condition (congruent condition, incongruent condition, isolated facial expression)

LOAD: Perceptual load (low load, high load)

Table 15. ANOVA Results for Accuracy in Experiment 6A

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
FA	1	50.02	0.14	0.717	0.007
Within-subjects effects					
CON	1.26	32696.36	31.38	0.000	0.611
LOAD	1	45.03	0.14	0.714	0.007
Interaction effects					
*FA*CON*LOAD*	1.53	112.64	0.52	0.553	0.025
*FA*CON*	1.26	828.29	0.80	0.409	0.038
*FA*LOAD*	1	78.91	0.24	0.629	0.012
*CON*LOAD*	1.53	238.51	1.10	0.332	0.052

Note: FA: Facial expression (disgust, fear)

CON: Consistency condition (congruent condition, incongruent condition, isolated facial expression)

LOAD: Perceptual load (low load, high load)

Table 16. ANOVA Results for Reaction Times in Experiment 6A

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
FA	1	171513.88	1.05	0.317	0.050
Within-subjects effects					
CON	1.49	27470.82	1.56	0.223	0.072
LOAD	1	1106427.76	18.15	0.000	0.476
Interaction effects					
*FA*CON*LOAD*	2	696.85	0.08	0.926	0.004
*FA*CON*	1.49	46642.79	1.96	0.167	0.089
*FA*LOAD*	1	67362.00	1.11	0.306	0.052
*CON*LOAD*	2	8098.72	0.90	0.415	0.043

Note: FA: Facial expression (disgust, fear)

CON: Consistency condition (congruent condition, incongruent condition, isolated facial expression)

LOAD: Perceptual load (low load, high load)

Table 17. ANOVA Results for d' in Experiment 6B

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
FA	1	0.05	0.05	0.832	0.002
Within-subjects effects					
CON	2	0.21	0.77	0.470	0.037
LOAD	1	25.86	40.69	0.000	0.670
Interaction effects					
*FA*CON*LOAD*	2	0.18	0.56	0.577	0.027
*FA*CON*	2	0.50	1.79	0.181	0.082
*FA*LOAD*	1	3.42	5.39	0.031	0.210
*CON*LOAD*	2	0.74	2.29	0.114	0.103

Note: FA: Facial expression (disgust, fear)

CON: Consistency condition (congruent condition, incongruent condition, isolated facial expression)

LOAD: Perceptual load (low load, high load)

Table 18. ANOVA Results for Accuracy in Experiment 6B

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
FA	1	2934.42	6.58	0.019	0.257
Within-subjects effects					
CON	1.30	22517.55	26.98	0.000	0.587
LOAD	1	2566.90	11.36	0.003	0.374
Interaction effects					
*FA*CON*LOAD*	2	26.90	0.15	0.859	0.008
*FA*CON*	1.30	890.09	1.07	0.332	0.053
*FA*LOAD*	1	11.68	0.49	0.491	0.025
*CON*LOAD*	2	191.67	1.09	0.346	0.054

Note: FA: Facial expression (disgust, fear)

CON: Consistency condition (congruent condition, incongruent condition, isolated facial expression)

LOAD: Perceptual load (low load, high load)

Table 19. ANOVA Results for Reaction Times in Experiment 6B

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
FA	1	44645.12	0.21	0.652	0.011
Within-subjects effects					
CON	1.47	206540.14	9.62	0.002	0.336
LOAD	1	1358729.30	29.25	0.000	0.606
Interaction effects					
*FA*CON*LOAD*	2	20646.72	2.45	0.100	0.114
*FA*CON*	1.47	10229.92	0.48	0.568	0.024
*FA*LOAD*	1	27341.93	0.59	0.452	0.030
*CON*LOAD*	2	20464.72	2.45	0.100	0.114

Note: FA: Facial expression (disgust, fear)

CON: Consistency condition (congruent condition, incongruent condition, isolated facial expression)

LOAD: Perceptual load (low load, high load)

Table 20. ANOVA Results for Log Transformation of IE scores in Experiment 6B

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
FA	1	0.26	1.29	0.270	0.064
Within-subjects effects					
CON	1.07	5.28	20.46	0.000	0.519
LOAD	1	0.74	25.68	0.000	0.575
Interaction effects					
*FA*CON*LOAD*	1.21	0.01	0.14	0.759	0.007
*FA*CON*	1.07	0.00	0.00	0.980	0.000
*FA*LOAD*	1	0.01	0.21	0.653	0.011
*CON*LOAD*	1.21	0.00	0.06	0.854	0.003

Note: FA: Facial expression (disgust, fear)

CON: Consistency condition (congruent condition, incongruent condition, isolated facial expression)

LOAD: Perceptual load (low load, high load)

Table 21. ANOVA Results for d' combining Experiment 6A and 6B

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
FA	1	0.2	0.17	0.681	0.004
MOD	1	24.86	21.64	0	0.357
Within-subjects effects					
CON	2	0.24	0.84	0.434	0.021
LOAD	1	62.09	98.64	0	0.717
Interaction effects					
*FA*MOD*CON*LOAD*	2	0.29	0.99	0.376	0.025
*CON*FA*MOD*	2	0.09	0.32	0.726	0.008
*LOAD*FA*MOD*	1	0.76	1.2	0.28	0.03
*CON*LOAD*FA*	2	0.03	0.11	0.896	0.003
*CON*LOAD*MOD*	2	0.65	2.23	0.114	0.054
*CON*FA*	2	0.5	1.78	0.176	0.044
*CON*MOD*	2	0.36	1.28	0.283	0.032
*LOAD*FA*	1	3.45	5.48	0.024	0.123
*LOAD*MOD*	1	0.76	1.2	0.28	0.03
*FA*MOD*	1	0.14	0.13	0.725	0.003
*CON*LOAD*	2	0.29	0.97	0.382	0.024

Note: FA: Facial expression (disgust, fear)

MOD: Modality (visual, auditory)

CON: Consistency condition (congruent condition, incongruent condition, isolated facial expression)

LOAD: Perceptual load (low load, high load)

Table 22. ANOVA Results for Log Transformation of IE scores combining Experiment 6A and 6B

Source	df	Mean Square	F	Sig	Partial Eta Squared
Between-subjects effects					
FA	1	0.04	0.27	0.604	0.007
MOD	1	0.13	0.9	0.349	0.022
Within-subjects effects					
CON	1.15	10.33	45.85	0	0.54
LOAD	1	0.55	10.12	0.003	0.206
Interaction effects					
*FA*MOD*CON*LOAD*	1.25	0.01	0.14	0.767	0.004
*CON*FA*MOD*	1.15	0.05	0.21	0.68	0.005
*LOAD*FA*MOD*	1	0.03	0.47	0.498	0.012
*CON*LOAD*FA*	1.25	0	0.05	0.87	0.001
*CON*LOAD*MOD*	1.25	0.08	1.04	0.331	0.026
*CON*FA*	1.15	0	0.18	0.713	0.004
*CON*MOD*	1.15	0.04	0.19	0.698	0.005
*LOAD*FA*	1	0	0.04	0.837	0.001
*LOAD*MOD*	1	0.24	4.4	0.042	0.101
*FA*MOD*	1	0.27	1.83	0.183	0.045
*CON*LOAD*	1.25	0.1	1.33	0.263	0.033

Note: FA: Facial expression (disgust, fear)
MOD: Modality (visual, auditory)
CON: Consistency condition (congruent condition, incongruent condition, isolated facial expression)
LOAD: Perceptual load (low load, high load)

Appendix B Stimuli Used in Current Experiments

Figure 1. Body Stimuli Evaluated in Experiment A.

Note: All images in Appendix B fit the pages. The real size of images used in experiments is described in specific chapter.

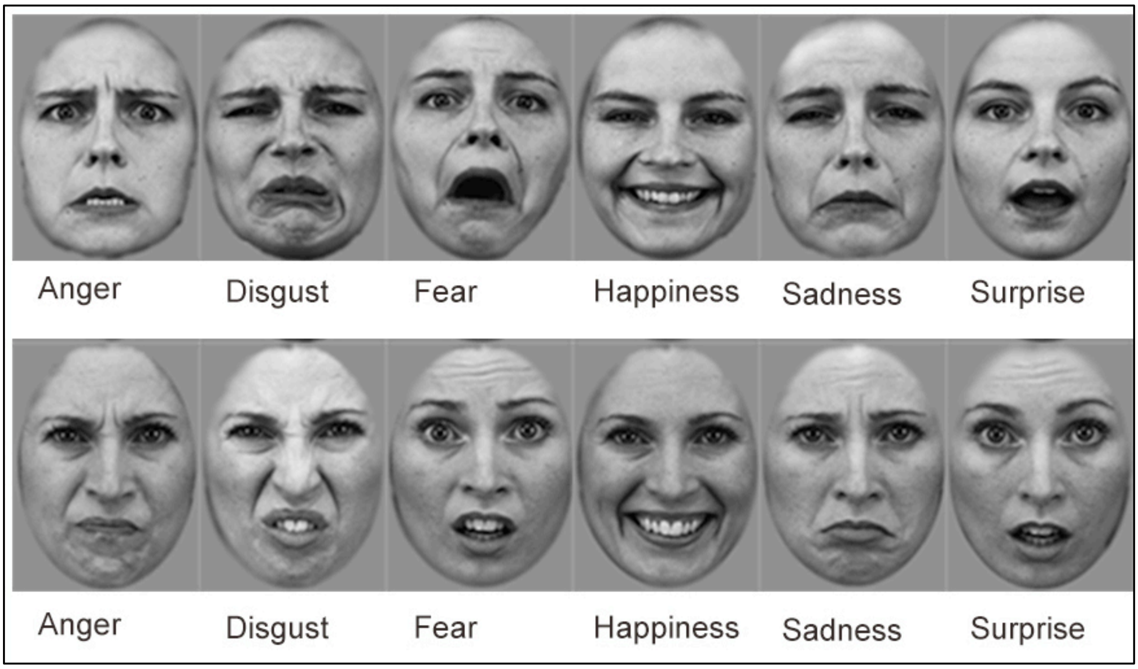

Figure 2. Isolated Female Facial Expressions in Experiment 1 and 2.

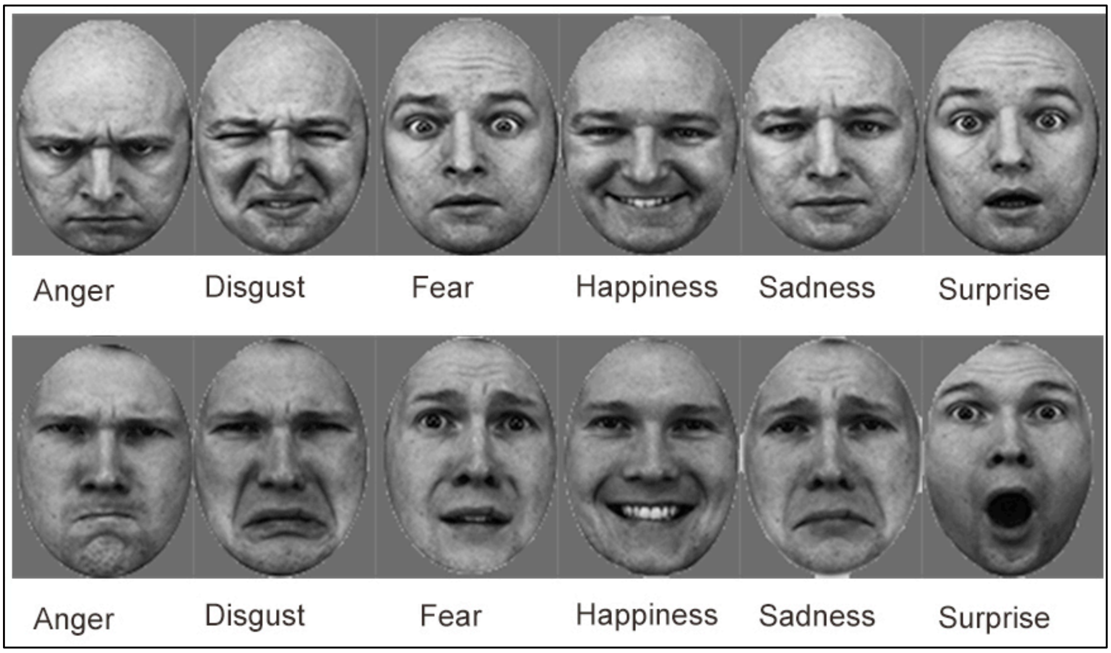

Figure 3. Isolated Male Facial Expressions in Experiment 1 and 2.

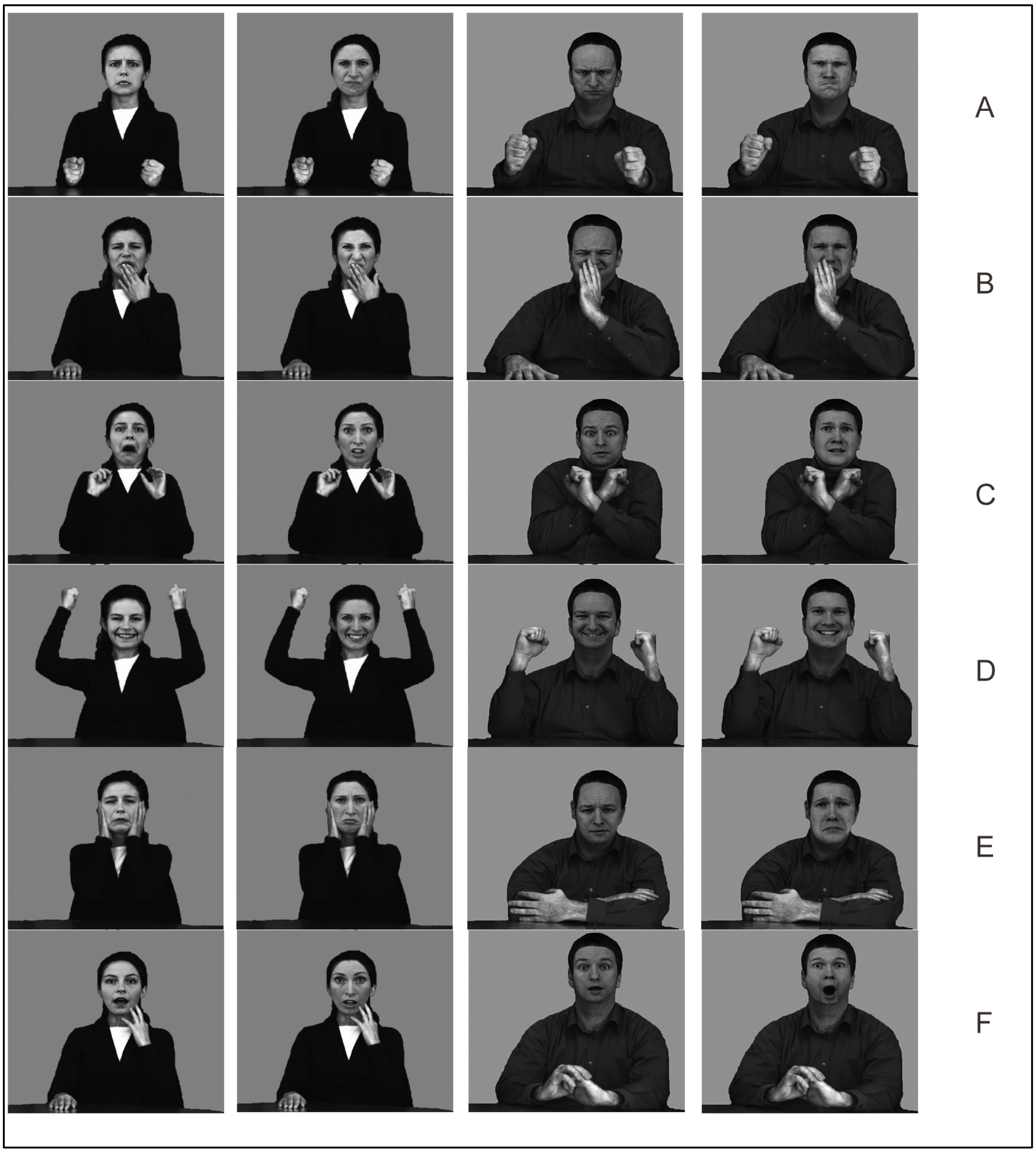

Figure 4. Congruent Face-body Stimuli in Experiment 1. The images in row A are anger face - anger body stimuli. The images in row B are disgust face - disgust body stimuli. The images in row C are fear face - fear body stimuli. The images in row D are happiness face – happiness body stimuli. The images in row E are sadness face – sadness body stimuli. The images in row F are surprise face – surprise body stimuli.

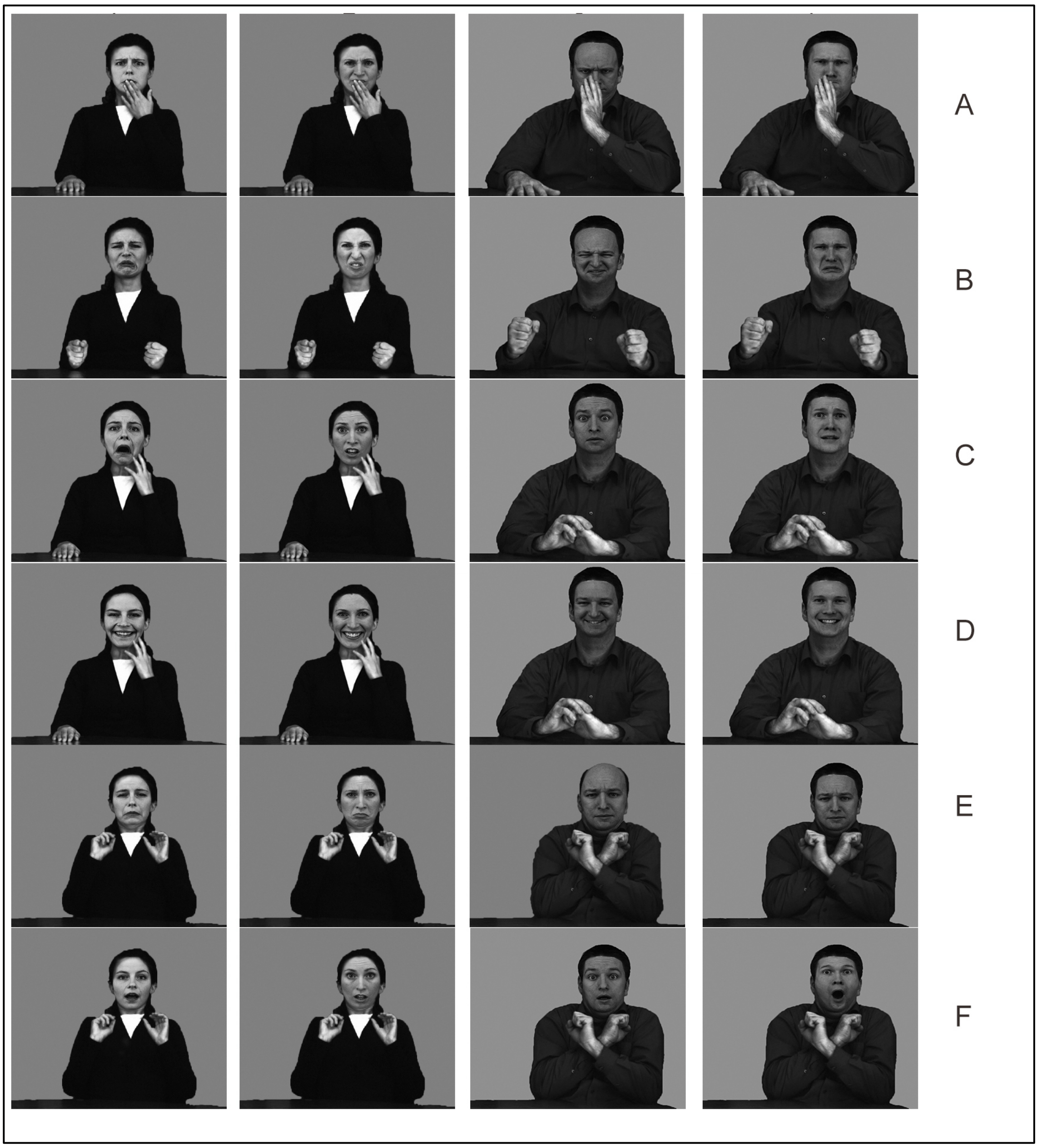

Figure 5. High Similarity Face-body Stimuli in Experiment 1. The images in row A are anger face - disgust body stimuli. The images in row B are disgust face - anger body stimuli. The images in row C are fear face - surprise body stimuli. The images in row D are happiness face – surprise body stimuli. The images in row E are sadness face – fear body stimuli. The images in row F are surprise face – fear body stimuli.

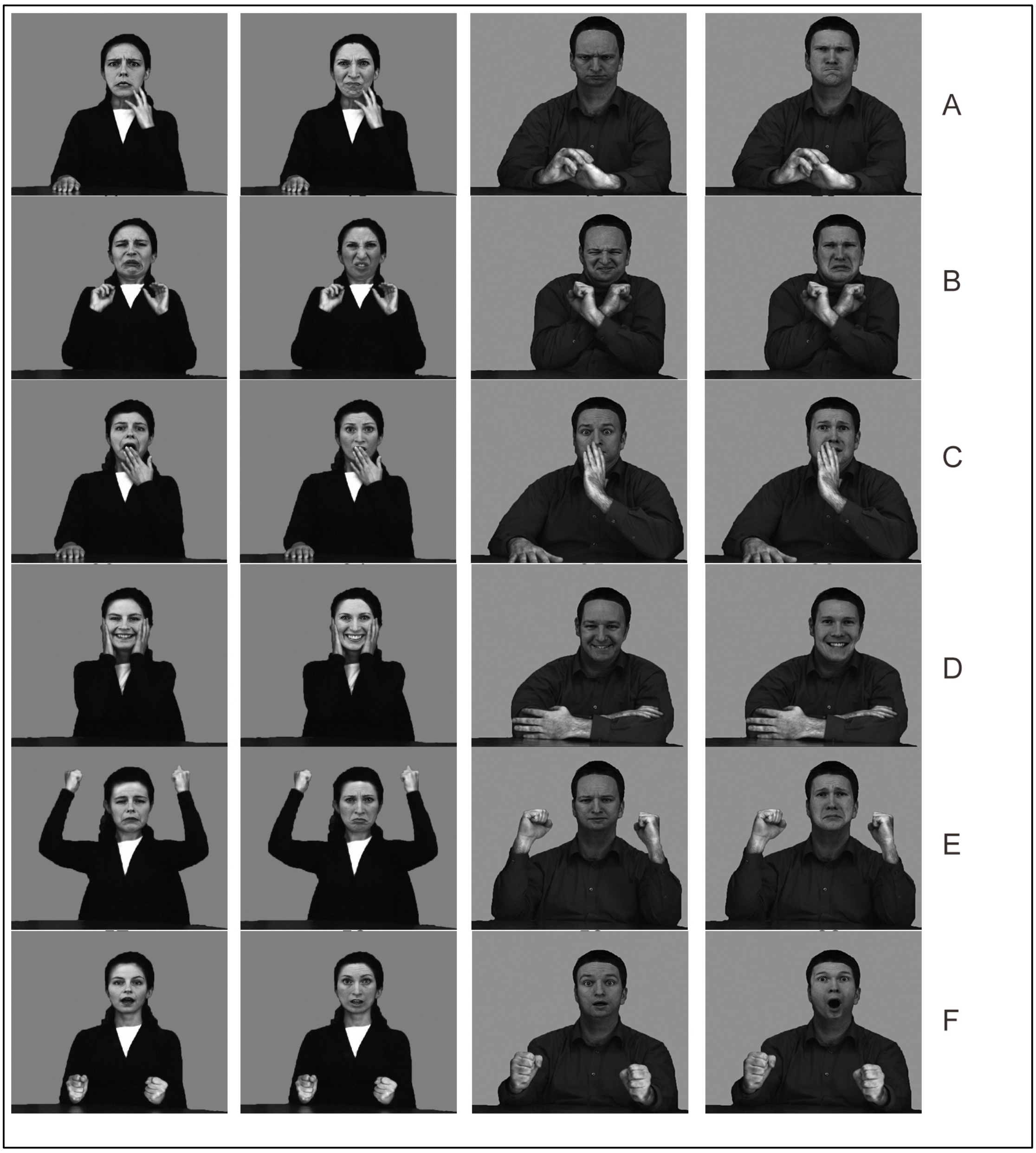

Figure 6. Low Similarity Face-body Stimuli in Experiment 1. The images in row A are anger face – surprise body stimuli. The images in row B are disgust face - fear body stimuli. The images in row C are fear face - disgust body stimuli. The images in row D are happiness face – sadness body stimuli. The images in row E are sadness face – happiness body stimuli. The images in row F are surprise face – anger body stimuli.

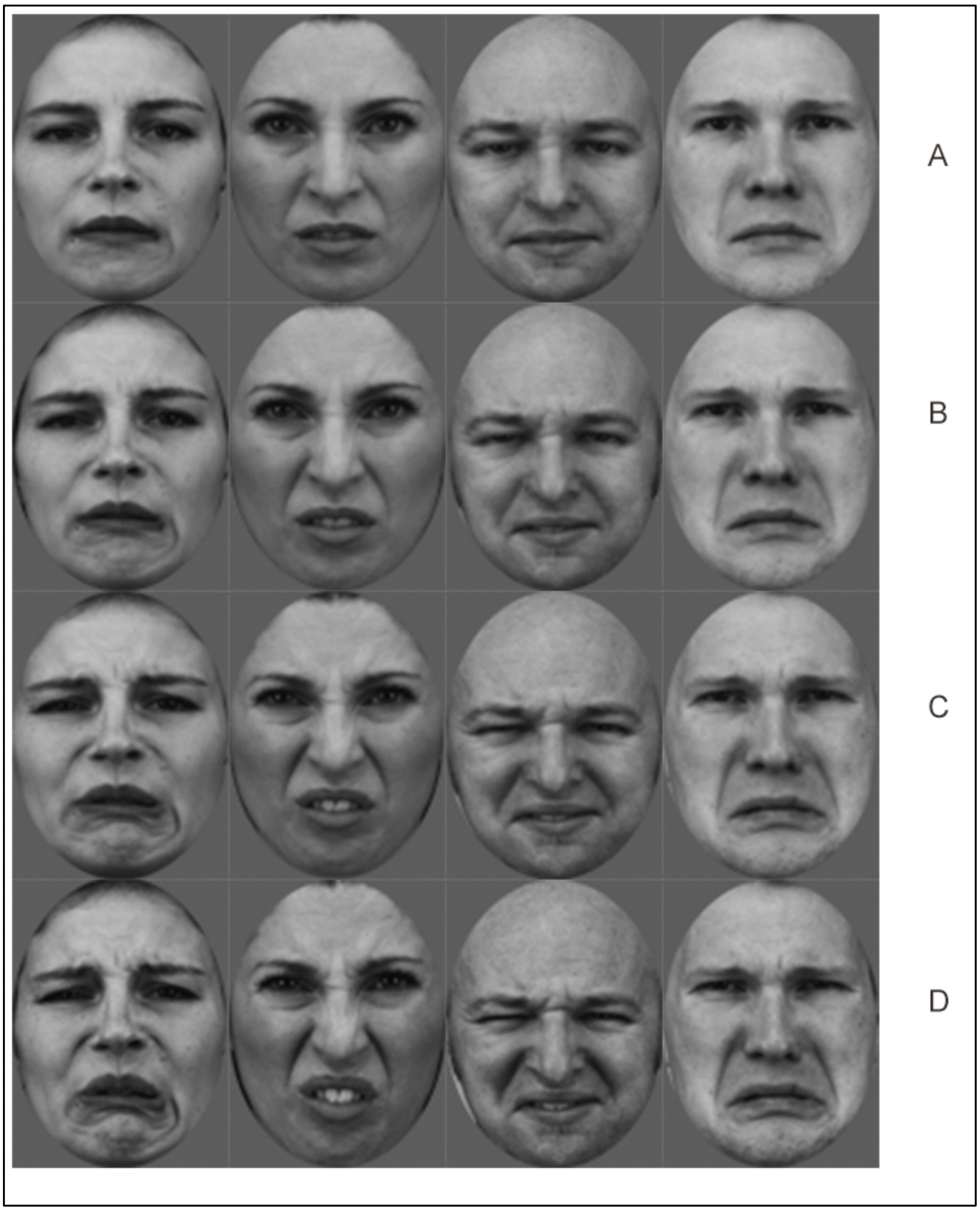

Figure 7. Isolated Disgust Facial Expressions in Experiment 3 and 4. The images in row A are disgust facial expressions at the level of 40% intensity. The images in row B are disgust facial expressions at the level of 60% intensity. The images in row C are disgust facial expressions at the level of 80% intensity. The images in row D are disgust facial expressions at the level of 100% intensity.

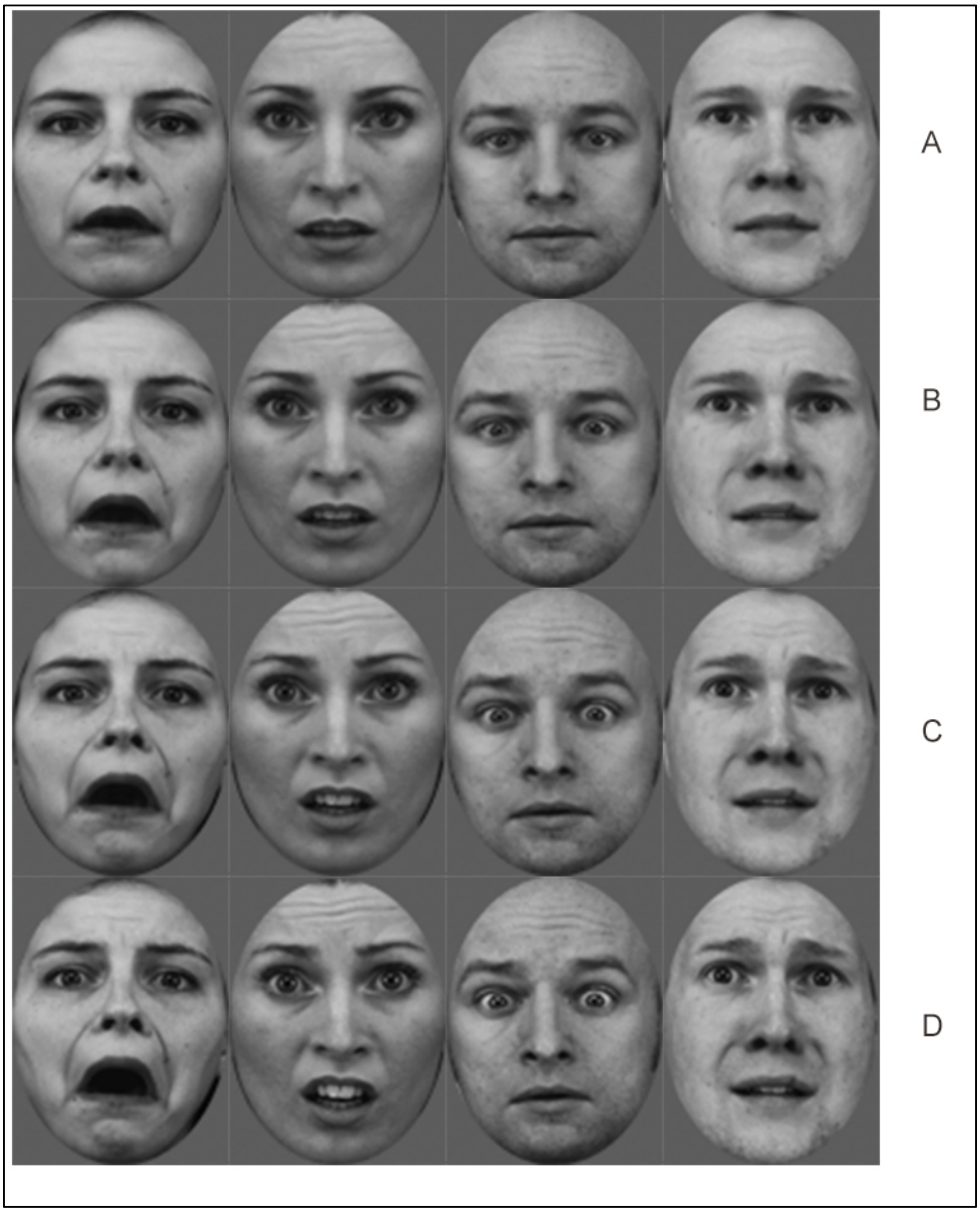

Figure 8. Isolated Fear Facial Expressions in Experiment 3 and 4. The images in row A are fear facial expressions at the level of 40% intensity. The images in row B are fear facial expressions at the level of 60% intensity. The images in row C are fear facial expressions at the level of 80% intensity. The images in row D are fear facial expressions at the level of 100% intensity.

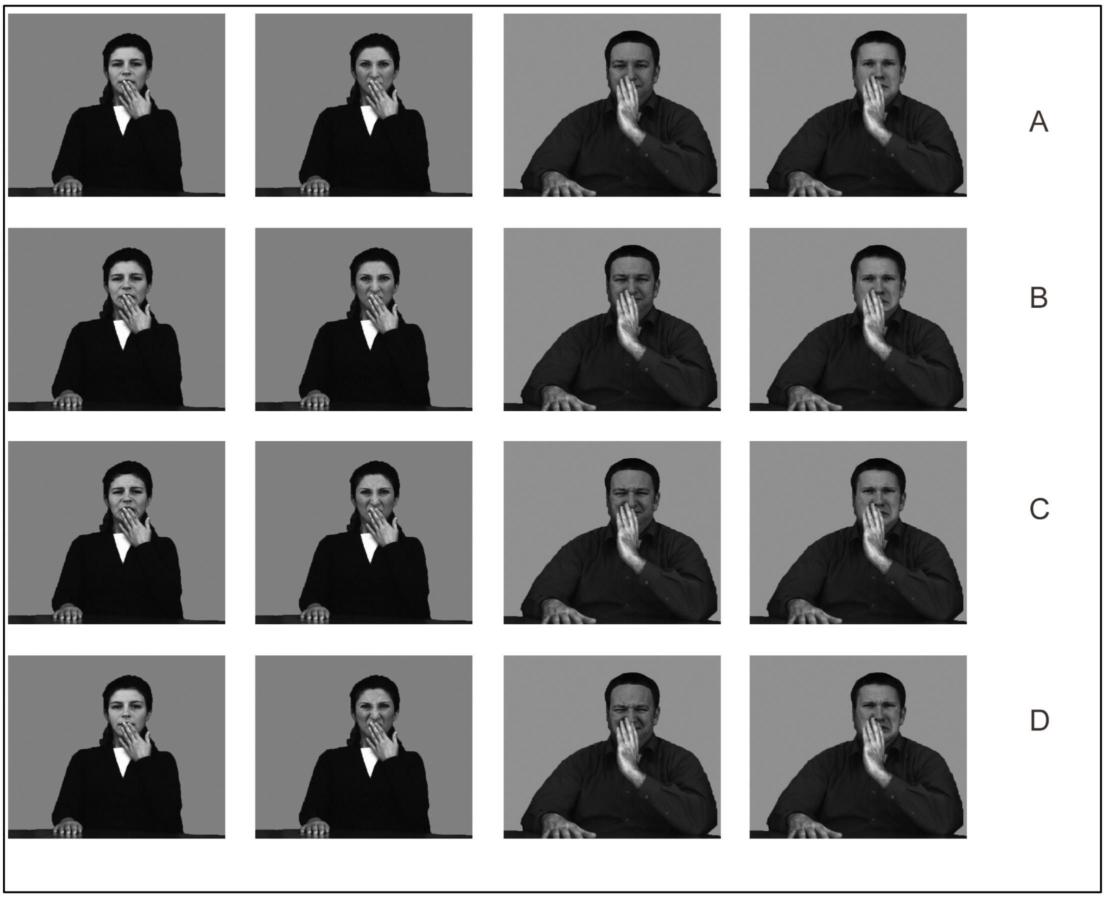

Figure 9. Disgust Face – Disgust Body (Congruent) Stimuli in Experiment 3. The images in row A are disgust face – disgust body stimuli at the level of 40% intensity. The images in row B are disgust face – disgust body stimuli at the level of 60% intensity. The images in row C are disgust face – disgust body stimuli at the level of 80% intensity. The images in row D are disgust face – disgust body stimuli at the level of 100% intensity.

Figure 10. Fear Face – Fear Body (Congruent) Stimuli in Experiment 3. The images in row A are fear face – fear body stimuli at the level of 40% intensity. The images in row B are fear face – fear body stimuli at the level of 60% intensity. The images in row C are fear face – fear body stimuli at the level of 80% intensity. The images in row D are disgust face – disgust body stimuli at the level of 100% intensity.

Figure 11. Disgust Face – Anger Body (Incongruent) Stimuli in Experiment 3. The images in row A are disgust face – anger body stimuli at the level of 40% intensity. The images in row B are disgust face – anger body stimuli at the level of 60% intensity. The images in row C are disgust face – anger body stimuli at the level of 80% intensity. The images in row D are disgust face – anger body stimuli at the level of 100% intensity.

Figure 12. Fear Face – Surprise Body (Incongruent) Stimuli in Experiment 3. The images in row A are fear face – surprise body stimuli at the level of 40% intensity. The images in row B are fear face – surprise body stimuli at the level of 60% intensity. The images in row C are fear face – surprise body stimuli at the level of 80% intensity. The images in row D are disgust face – surprise body stimuli at the level of 100% intensity.

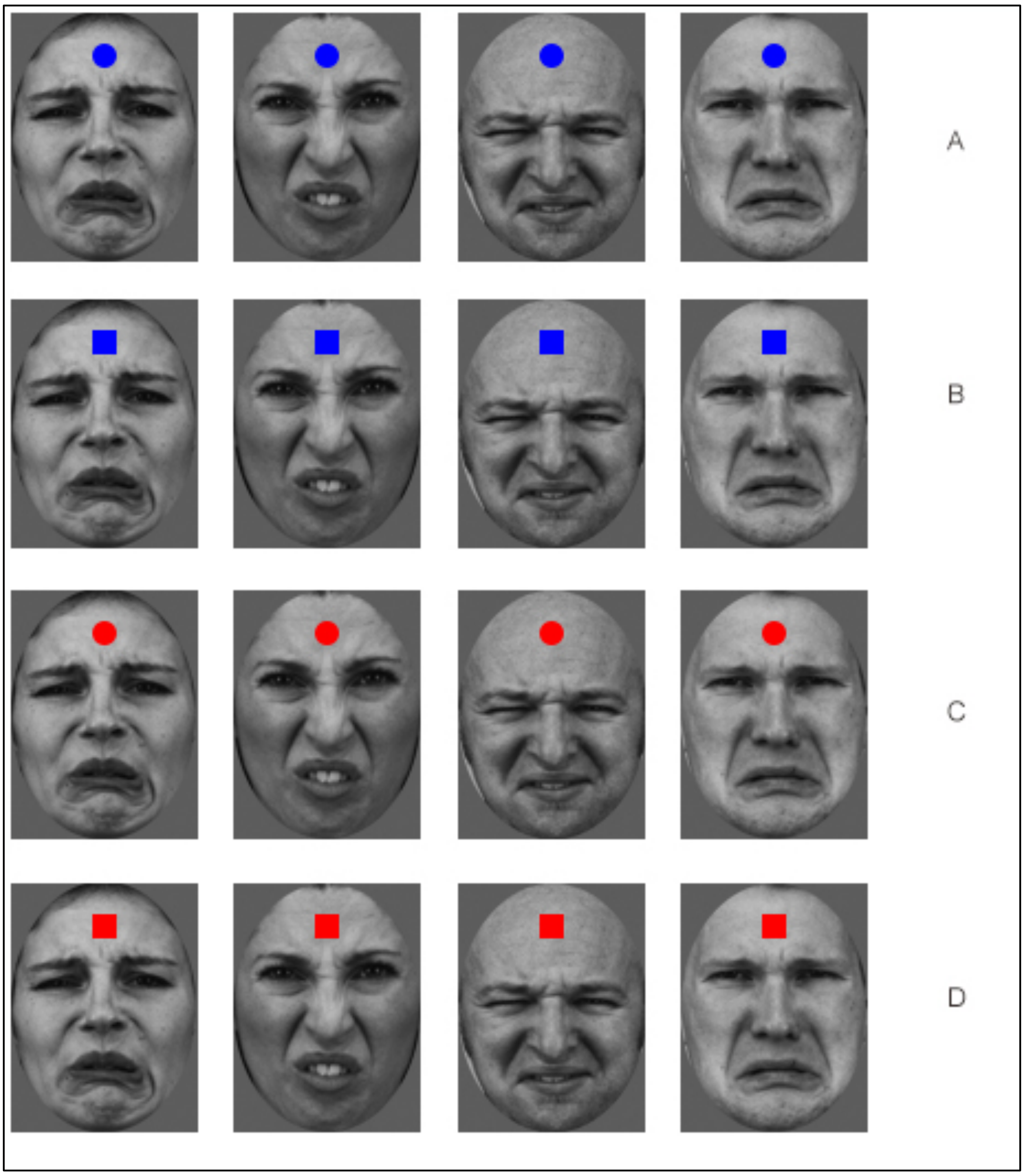

Figure 13. Isolated Disgust Faces in Experiment 5 and 6. The images in row A are disgust faces with a blue circle on the centre of forehead. The images in row B are disgust faces with a blue square on the centre of forehead. The images in row C are disgust faces with a red circle on the centre of forehead. The images in row D are disgust faces with a red square on the centre of forehead.

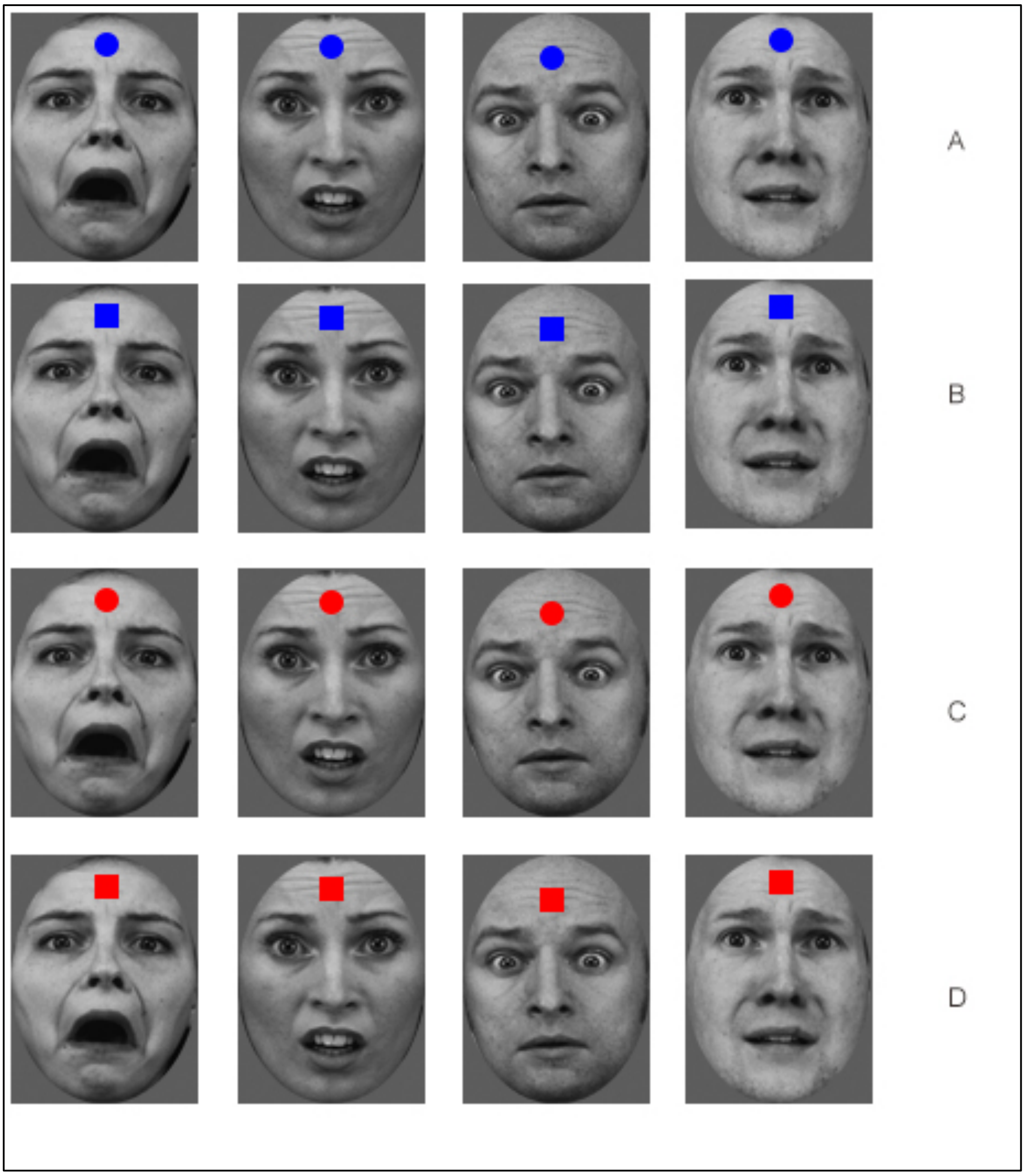

Figure 14. Isolated Fear Faces in Experiment 5 and 6. The images in row A are fear faces with a blue circle on the centre of forehead. The images in row B are fear faces with a blue square on the centre of forehead. The images in row C are fear faces with a red circle on the centre of forehead. The images in row D are fear faces with a red square on the centre of forehead.

Figure 15. Disgust Face – Disgust Body (Congruent) Stimuli in Experiment 5. The images in row A are disgust face – disgust body stimuli with a blue circle on the centre of forehead. The images in row B are disgust face – disgust body stimuli with a blue square on the centre of forehead. The images in row C are disgust face – disgust body stimuli with a red circle on the centre of forehead. The images in row D are disgust face – disgust body stimuli with a red square on the centre of forehead.

Figure 16. Fear Face – Fear Body (Congruent) Stimuli in Experiment 5. The images in row A are fear face – fear body stimuli with a blue circle on the centre of forehead. The images in row B are fear face – fear body stimuli with a blue square on the centre of forehead. The images in row C are fear face – fear body stimuli with a red circle on the centre of forehead. The images in row D are fear face – fear body stimuli with a red square on the centre of forehead.

Figure 17. Disgust Face – Anger Body (Incongruent) Stimuli in Experiment 5. The images in row A are disgust face – anger body stimuli with a blue circle on the centre of forehead. The images in row B are disgust face – anger body stimuli with a blue square on the centre of forehead. The images in row C are disgust face – anger body stimuli with a red circle on the centre of forehead. The images in row D are disgust face – anger body stimuli with a red square on the centre of forehead.

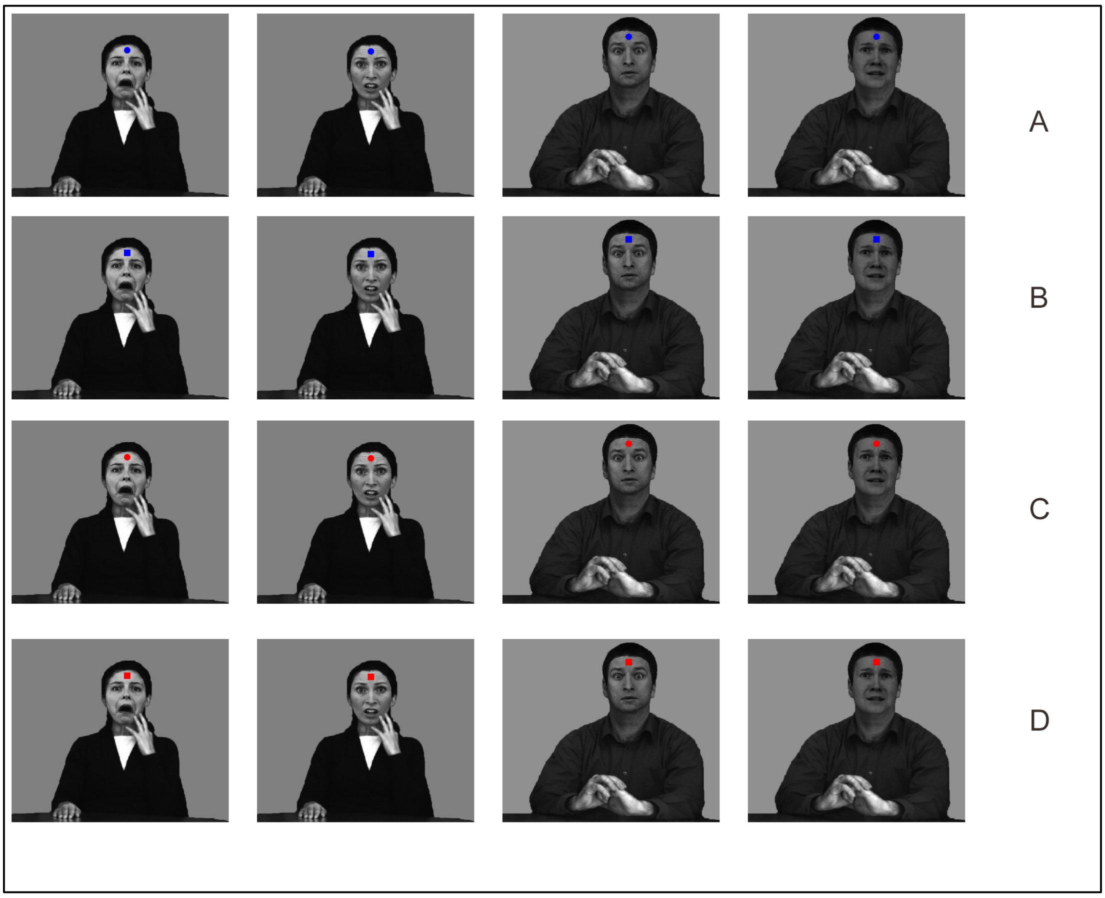

Figure 18. Fear Face – Surprise Body (Incongruent) Stimuli in Experiment 5. The images in row A are fear face – surprise body stimuli with a blue circle on the centre of forehead. The images in row B are fear face – surprise body stimuli with a blue square on the centre of forehead. The images in row C are fear face – surprise body stimuli with a red circle on the centre of forehead. The images in row D are fear face – surprise body stimuli with a red square on the centre of forehead.