國立臺灣大學理學院心理學研究所

博士論文

Graduate Institute of Psychology

College of Science

National Taiwan University

Doctoral Dissertation

持續閃現抑制下的無意識情緒訊息處理

Unconscious Processing of Emotional Information under Continuous Flash Suppression

楊詠皓

Yung-Hao Yang

指導教授：葉素玲 博士

Advisor: Su-Ling Yeh, Ph.D.

中華民國 105 年 1 月

January, 2016

致謝

前兩天的時候台北下了雪，今天的太陽卻又暖和地了得，差不多是在這樣一個奇妙氣候的冬天裡完成人生的一個階段。從碩士班進入實驗室至今，算算也將近十個年頭。十年，是個不算長也不算短的日子。這篇論文得以完成，一路上要感謝的人很多。

首先要感謝我的指導教授葉素玲老師，允許我在這些年的漫天想像中不斷嘗試各種可能，並以無比的耐心修正我總是渙散的思考。在我心裡，一直保有一個畫面。那是約莫五六年前，實驗室在歐洲視覺年會前，順道去了趟德國最高的楚格峰(Mt. Zugspitze)。那時山嵐瀰漫峰谷之間，葉老師在延綿彎曲的霧途中，時而在前方引領我們，時而在後方提醒著我們。那趟旅程就像這幾年的縮影，老師的言教與身教也如同實驗室的名稱─Explorer of Perception and Attention 一般，至始至終地提醒我以股勇敢而謹慎的姿態去探索科學中各種未知。

感謝在論文計畫以及論文最後階段的口試委員們，謝謝阮啟弘老師，李金鈴老師，許儷絹老師，黃從仁老師，黃淑麗老師，黃榮村老師，龔充文老師。謝謝您們花了許多時間仔細地審視我的研究，從各種角度提供全面的觀點，提點我論文的細節與定位，使這份論文最終得以更完整的樣貌呈現。也感謝在德國與日本做訪問學生時的指導老師，Dr. Andreas Bartels 與程康老師提供學習的機會。這些寶貴的研究經驗培養我進行獨立研究的能力。

謝謝實驗室的學長姐與同學們。謝謝山源學長與心怡學姐從我一進入實驗室就耐心地進行傳承，我依舊懷念那段大家固定在家中的聚會。謝謝彥文，歡偉，詠瑄，魁安，柏寬在撰寫論文這段期間的協助與支持，讓我在最後的日子裡能不顧一切地全力衝刺。謝謝那些參與我實驗的人們，因為有你們的參與，讓我們得以有機會增進對人類知覺處理的了解。謝謝攝影社的一切，是你們讓我見識到學以致用的花火。特別感謝亦師亦友的菘瑋，還記得有次花了整天時間教我把程式寫得簡潔而到位。謝謝社上提供這樣的場域，在暗房的時光總能讓我沉澱所有紊亂。

最後，謝謝我的父母，這段路雖然走的緩慢，卻以無與倫比的耐心與鼓勵，支持我朝目標前進。謝謝我的阿嬤，每天打電話回家時總是叮嚀我要照顧身體，求菩

薩保佑我一切平安。謝謝大漢與軒田姊夫，不厭其煩的提供生活與學術上的各種協助。謝謝詠惠，總是體貼地幫我留意新出的書和處理各種雜事。

最後的最後，謝謝欣惠，總在我忘記吃飯的時候叫我吃飯，陪我一起熬夜寫論文。謝謝妳一直以來的一切擔待與陪伴，由於妳的默默支持與包容，這份論文得以完成。

摘要

意識的科學研究在近年來受到廣泛的注意，而了解無意識處理的限度是理解意識本質的重要步驟。由於持續閃現抑制派典能使各種訊息進行長時間且維持注意力的呈現，近來成為探討無意識處理的傑出工具。然而，沿著視覺腹側路徑的訊息能進行何種程度的無意識處理仍有很大的爭議。為了回答這個議題，在此論文中的三個研究中我們藉由不同的測量方式檢驗臉部表情的情緒內容可否在持續閃現抑制下被處理。研究一顯示我族熟悉性而非方向熟悉性可以調節不同臉部表情在持續閃現抑制下偵測的時間。研究二顯示臉部表情在持續閃現下的偵測時間以及眼動資料，皆會受到與其意涵一致的情緒聲音所促進。研究三顯示持續閃現抑制下的情緒促發物能影響後續目標的情緒判斷。這些聚合的研究結果支持臉部表情—此類沿著視覺腹側路徑的訊息—可以被無意識處理。此外，上述結果促使我們結合不同測量方法的優點，進而提出新的實驗派典—凝視同步持續閃現抑制派典，預期此派典將作為未來探討無意識處理的有力方法。

關鍵詞：無意識處理、情緒訊息、臉部表情、持續閃現抑制

Unconscious Processing of Emotional Information under Continuous Flash Suppression

Yung-Hao Yang

Abstract

Scientific study of consciousness has gained great attention in recent decades. Knowing the capacity of unconscious processing is an important step to understand the nature of consciousness. The recent development of Continuous flash suppression (CFS) paradigm is an excellent tool to probe the unconscious processing of various kinds of information with full processing time and attention. However, it is still hotly debated about which level of information along the visual-ventral pathway can be processed under CFS. We approach this issue in the current thesis by conducting three studies with different measurements to test whether emotional content of facial expression can be processed under CFS. Study 1 showed that own-race familiarity but not orientation familiarity can modulate the time of different facial expression for breaking CFS. Study 2 showed that affective voice with coherent valence can facilitate the processing of corresponding facial expression under CFS, which was supported by breaking-CFS time and eye-movements data. Study 3 showed that invisible affective prime under CFS can affect emotional judgment of subsequent target. These convergent results support that facial expression—which is supposedly processed along ventral pathway—can be processed unconsciously. Moreover, above findings enlighten us to propose a new paradigm, the gaze–contingent CFS paradigm, which combines several advantages from different measurements. We expect this paradigm that may provide a powerful method to probe unconscious processing in the future

Keywords: unconscious processing, emotional information, facial expression, continuous flash suppression

Table of contents

Introduction

Understanding the nature of consciousness with scientific approach has gained great attention in recent decades (Koch & Crick, 1990), and significant progress on relevant issues about consciousness has been made from interdisciplinary studies across different fields including philosophy (Baars, 2005; Block, 2005; Chalmers, 1998), psychology (Brigard & Prinz, 2010; Tsuchiya & Koch, 2008; van Boxtel, Tsuchiya, & Koch, 2010; Van den Bussche, Hughes, Humbeeck, & Reynvoet, 2010), and neuroscience (Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006; Koch & Tsuchiya, 2007; Lamme, 2003, 2004).

According to Chalmers (1998), the scientific study of consciousness can be divided into two questions: the easy problem and the hard problem. The *easy problem* is relevant to the explanations of how we perceive, integrate, and react to the world, which is relatively "easy" to solve because it can be answered by objective measurements with current scientific approaches. The h*ard problem*, however, is hard to explain by objective measurement because it is relevant to the explanations of subjective experience such as "what is it like to be a bat?" (Nagel, 1974), the redness of red, and the sweetness of sweet, etc.. Based on the general assumption of scientific studies of consciousness that subjective experience also arises from neural systems, it becomes hard to answer the question that how a physical brain can generate the non-physical entity of phenomenal experience. Therefore, there is an "explanation gap" between the easy problem and the hard problem, with the former being able to be answered by current scientific measures and the latter that cannot be explained. Consequently, nowadays most researchers focus on the "easy" problem and tackle issues like which mental process and neural mechanism are involved during conscious percept versus unconscious percept (Kouider, 2009).

Koch and Crick (2001) proposed an idea called the neural correlate of consciousness (NCC), which refers to the minimal neural activities required to generate conscious percept. Focusing on the conscious processes alone cannot satisfy the condition of "minimum neural activities" because some processes can occur even without consciousness (i.e., unconscious processing). To study the NCC, it is important to distinguish neural activations that are involved during conscious processing from those during unconscious processing (Rees, Kreiman, & Koch, 2002). We approach this issue in the current thesis by asking the question that to what extent information, especially emotional information, can be processed unconsciously. By knowing the capacity of unconscious processing, it will enlighten our understanding of the nature of consciousness.

Understanding the Unconscious Processing with Scientific Approach

According to Laureys (2005), the definition of consciousness can be sorted and placed onto a two dimensional space. One is *level of arousal*, such as wakefulness, sleep, coma, etc., which refers to the *state* of consciousness. The other is *representational capacity*, such as seeing the light, hearing the sound, touching the texture, etc., which refers to the *access* of consciousness. This thesis focuses on the latter one; that is, whether people can process information without conscious of its representation. We adopt two operational definitions to evaluate this unconscious processing in this thesis (Kunimoto, Miller, & Pashler, 2001; Merikle & Daneman, 2000; Reingold & Merikle, 1988) and discuss the pros and cons of different paradigms that probe unconscious processing.

The operational definition of unconscious processing

Subjective measurement. The most institutive way to define consciousness is based on participants' subjective reports. This introspective measurement assumes that only participants themselves can access their conscious experience. Therefore, whether a process is defined as conscious or unconscious depends on an observer's reports of seeing or not. The advantage of this measurement is that the defined consciousness is closely linked to subjective experience, and thus it provides a first-person perspective from observer's own consciousness (Merikle, Smilek, & Eastwood, 2001). One drawback, however, is that this measurement may be heavily affected by response bias. For example, different participants may use different criteria to report their visual percept, and thus overestimate or underestimate their subjective consciousness.

Objective measurement. Another way to define consciousness is using an objective measurement. That is, participants are forced to detect or discriminate a to-be tested-stimulus, and if the performance is at chance level, then it suggests that participant is unconscious about the stimulus. The advantage of this measurement is that definition of consciousness is independent of subjective criteria, and can merely depends on objective performance (Kunimoto et al., 2001). However, defining consciousness based on this objective measurement has been criticized as being too liberal (Merikle et al., 2001). For example, in the case of blindsight patients, they would be defined as being conscious of the stimuli according to the objective criterion since their performance is above chance level in some task, even when they deny seeing the stimulus (L. Weiskrantz, 1996). One drawback about using the objective measurement may underestimate the capacity of unconscious processing because all the abilities reserved in blindsight patients would be treated as carried out in this objectively defined capacity of "consciousness".

Although subjective measurement and objective measurement of consciousness are based on different assumptions of assess consciousness, they nevertheless lead to similar conclusions (Merikle et al., 2001) for neurologically intact participants. For example, Del Cul, Dehaene, and Leboyer (2006) varied intervals between prime and mask to assess the conscious threshold in the backward-masking paradigm, and found that the conscious thresholds by both subjective and objective measurements were highly correlated ($r^2 = .96$, Figure 5 in Del Cul et al., 2006). This result implies that subjective criteria provide reliable measurement of consciousness as objective criteria. Thus, in this thesis, we used both subjective measurement (Study 1 and Study 2) and objective measurement (Study 3) to define unconscious processing.

Paradigms probing unconscious processing

There are several psychophysical paradigms that can render visual stimuli invisible and thus probe unconscious processing of stimuli; however, different paradigms have their own pros and cons (see Blake & Logothetis, 2002; Kim & Blake, 2005 for reviews). Basically, there are two types of Paradigms for unconscious processing. One is to distract *attention* away from visual stimuli, and the other is to reduce or interrupt *sensory input* (Kanai, Walsh, & Tseng, 2010; Merikle et al., 2001). Paradigms by distracting attention include attentional blink (Raymond, Shapiro, & Arnell, 1992) and inattentional blindness (Li & Yeh, 2007; Mack & Rock, 1998) in which by distracting attention away from the target, it cannot be perceived consciously, which can nevertheless be accessed with attention (Dehaene et al., 2006).

More relevant to this thesis is the paradigms interrupting sensory input. Previous studies have used varieties of such interrupting paradigms to probe unconscious processing. For example, in the commonly adopted masking paradigm, a mask that precedes or follows a briefly presented target can induce visually forward- or backward-masking so as to interrupt the incoming signal from being processed consciously. Other paradigms like visual crowding (Levi, 2008; Yeh, He, & Cavanagh, 2012), motion induced blindness (Bonneh, Cooperman, & Sagi, 2001; Hsu, Yeh, & Kramer, 2004, 2006), bistable figures (e.g., Young lady and old woman; Hsiao, Chen, Spence, & Yeh, 2012), and binocular rivalry (Alais & Blake, 2005; Y. C. Chen, Yeh, & Spence, 2011). Unlike visual masking and crowing that vary physical properties of stimulus to render information being processed unconsciously, in bistable figure and binocular rivalry participants' visual consciousness as to either one of the two percepts fluctuate when receiving the same visual stimulation. These paradigms are suitable to find the NCCs because neural events are related to the difference between unconscious and conscious perception without accompanied changes in stimuli. However, one obvious drawback is that conscious and unconscious percepts vary unpredictably (Blake & Logothetis, 2002; Kim & Blake, 2005).

Different manipulations probe different mechanisms of unconscious processing (Dehaene et al., 2006; Kanai et al., 2010), and it is important to choose a suitable paradigm for a particular purpose. For example, it is critical to distinguish between information that cannot be processed due to lack of attentional resource and information that can be processed unconsciously with attention (e.g., Watanabe et al., 2011). Since attention and consciousness are related but distinct (see Koch & Tsuchiya, 2012; van Boxtel et al., 2010 for reviews), adopting the paradigm of distracting attention would lead to the difficulty to dissociate the situations between that lacks attention and that lacks consciousness. In addition, information requires time to process, and it is also important to differentiate the situation in which unreportable information is due to insufficient processing time (e.g., Lo & Yeh, 2008) from the situation that it indeed cannot be processed unconsciously in nature. To better reveal the extent that information can be processed unconsciously with full processing time and attention, we adopted the continuous flash suppression (CFS; Jiang, Fang, Huang, & He, 2005; Tsuchiya & Koch, 2005).

The Continuous Flash Suppression Paradigm (CFS)

CFS is a variation of binocular rivalry; both implement the interocular suppression technique. In this paradigm, the stimulus is presented to one eye, and a train of high contrast dynamic masks is presented to the other eye. Because the masks consist of strong signals (i.e., high contrast and dynamic), the stimulus (usually weak in contrast) is interocularly suppressed by the masks. The processing of the suppressed stimulus is considered as unconscious processing.

There are several advantages by adopting this paradigm: (1) In the temporal domain, by manipulating contrast and spatial frequency between the stimulus and the masks (E. Yang & Blake, 2012), the stimulus can be suppressed from being consciously perceived for a relatively long duration (compared to masking and attentional blink) and the percept is relatively predictable (compared to motion induced blindness, bistable figure, and binocular rivalry). (2) In the spatial domain, the stimulus can be presented in any locations of the visual field (compared to crowding and motion induced blindness) and can adopt any kinds of stimuli (compared to bistable figure). (3) In the attentional domain, since this paradigm mainly interrupts the sensory input, the stimulus can be processed with full attention (compared to attentional blink and inattentional blindness). (4) In comparing the processing efficiency between different stimuli, this paradigm uses the same masks to render different stimuli invisible. It therefore can clearly differentiate the processing between different stimuli under the same suppressed situation (compared to the binocular rivalry, which compared the domain phase and suppressed phase between stimuli; Jiang, Costello, & He, 2007)

Processing capacity and limitation under CFS

To what extent information can be processed unconsciously? This is not only an important question for understanding the nature of consciousness, but also a critical issue for studies on interocular suppression. Using conventional binocular rivalry, several studies have demonstrated that low-level (Blake & Fox, 1974; Lehmkuhle & Fox, 1975; Wade & Wenderoth, 1978), but not high-level information (Sheinberg & Logothetis, 1997; Tong, Nakayama, Vaughan, & Kanwisher, 1998; Zimba & Blake, 1983) can be processed under interocular suppression. These results may be due to that dichoptic information is reciprocally competed with each other at early stages of visual processing (Blake & Logothetis, 2002; Sengpiel & Vorobyov, 2005), and residual high-level information is too weak to be processed under interocular suppression.

The recent development of CFS allows longer and predictable suppression time than conventional binocular rivalry (Tsuchiya, Koch, Gilroy, & Blake, 2006), and thus provides an excellent tool to assess whether high-level information can still be processed under interocular suppression (see Z. Lin & He, 2009 for a review). For example, Fang and He (2005) tested whether faces and tools revealed different processing under CFS by adopting brain imaging technique. The found higher activity in the dorsal pathway for tools than for faces, and weak activity in general in the ventral pathway. This finding supports two-streams theory that dorsal pathway is related to implicit guided action, and ventral pathway is related to explicit object recognition (Goodale & Milner, 1992; Milner & Goodale, 2008). Regarding the capacity and limitation of unconscious processing under CFS, we introduce below the findings based on the differentiation of dorsal and ventral pathways of visual system.

Information along the dorsal pathway. In line with the above finding (Jiang et al., 2005), several studies also showed that visual information under CFS can be processed through the dorsal pathway. For example, Almeida, Mahon, Nakayama, and Caramazza (2008) adopted pictures of tools and animals as stimuli with a category priming paradigm to test stimulus-specific effect under CFS. They found that priming effects were only found for tool target but not for animal target, supporting the idea that information along the dorsal pathway but not the ventral pathway can be processed under CFS (see also Almeida, Mahon, & Caramazza, 2010 for comparing the results with using stimuli of vehicles). Moreover, both grasping (Roseboom & Arnold, 2011) and proprioception (Salomon, Lim, Herbelin, Hesselmann, & Blanke, 2013) can be integrated with the corresponding invisible stimulus under CFS. For example, Salomon et al. (2013) asked participants to pose their hand with the palm up or down, and a masked hand picture with either congruent or incongruent posture as the real hand was presented under CFS. They found that the congruent masked hand picture released from suppression more quickly than the incongruent one. This result implies that not only visual modality but also cross-modal information along the dorsal pathway can be unconsciously integrated under CFS. In addition to the unconscious processing of action-relevant information, numerosity that is processed in the intra-parietal sulcus (IPS) can also be processed under CFS (Bahrami et al., 2010; Sklar et al., 2012). For instance, Bahrami et al. (2010) found that numerical primes in either a symbolic form (e.g., Arabic numerals) and non-symbolic form (e.g., dot patterns) under CFS can facilitate subsequent target enumeration (i.e., numerical-priming effect), and this effect was modulated by the numerical distance between the prime and the target. In summary, information processed along the dorsal pathway, such as action-relevant stimuli (tools, hand gesture, etc.) and numerical stimuli can be processed under interocular suppression.

Information along the ventral pathway. Relative to the dorsal pathway, it is still hotly debated about which level of information along the ventral pathway can be processed under interocular suppression. For example, E. Yang, Zald, and Blake (2007) measured the time that different facial expressions released from interocular suppression, and found that fearful faces had faster release time than neutral and happy faces. However, this result was also found when the faces were inverted or when only eye-white regions were presented. This observation begs the question that the result may have been caused by low-level properties, such as the high-contrast of large eye-white regions embedded in a fearful face. To address this issue, Gray, Adams, Hedger, Newton, and Garner (2013) created a particular kind of control faces by manipulating both face inversion and luminance polarity to interrupt emotional meaning but kept low-level properties as same as normal faces. The same result that fearful face gained preferential access was found for both normal faces and control faces, and thus they concluded that the seemingly unconscious processing of facial expression under CFS can be entirely explained by low-level properties. T. Stein and Sterzer (2012) reached a similar conclusion (but with different results) by using schematic face. In their study, they used schematic face to represent happy or angry emotion, and manipulated face in normal viewing, inverted viewing, without nose, and even without eyes (thus unlike a face). They consistently found that happy face released from suppression more quickly than angry face in all of the above conditions. Since happy face has parallel curvatures of mouth and face contour, they therefore concluded that this effect of happy face preferential access to consciousness was mainly related to figure-ground segregation but not emotional valence. These results suggest that although stimuli with different facial expressions can be processed under CFS, it remains unclear which level of information had been probed in previous studies.

Another way to test whether facial expression can be processed under interocular suppression is to test aftereffect generated by invisible face. For example, Adams, Gray, Garner, and Graf (2010) adopted different facial expressions under CFS, and found that emotional percept of the subsequent face stimulus can be changed by preceding invisible facial expression (i.e., facial-expression aftereffect). However, E. Yang, Hong, and Blake

(2010) argued against their result (Adams et al., 2010) that it might be due to the visible and invisible adapting faces interleaved within the same session, causing residual adaption from the visible trial to transfer to the invisible trial. Indeed, while the visibility of adapting face was manipulated in separate sessions, the facial-expression aftereffect disappeared when the adapting face was fully suppressed (E. Yang, Hong, et al., 2010).

These contradictory results leave unsolved the question whether facial expression, which is suggested to be processed along the ventral pathway, can be processed unconsciously under CFS. In this thesis, we conducted three studies with different approaches to address this issue.

Measurement under CFS

Different types of measurement have been designed to reveal unconscious processing under CFS, and these are also relevant to direct or indirect measurement based on subjectively and/or objectively defined unconscious processing. Each measurement has its pros and cons, and we summarize them below:

Breaking-CFS. The breaking-CFS is a direct measure with subjective criterion of unconscious processing. This paradigm measures the time a stimulus requires for breaking suppression by asking participants to detect the stimulus. The basic assumption is that unconscious processing of different stimuli may have different processing efficiency, and thus needs different suppression times to break suppression. Since the suppression time is measured as the latency between stimulus onset and the time that the stimulus is detected, the latency before target detection can be an index of unconscious processing.

Pros. (1) Since the breaking-CFS is a direct measure, almost any kinds of stimulus can be manipulated directly to test the efficiency under CFS (compared to the aftereffect-CFS, see below). (2) There are some individual differences for successful suppression of

CFS (cf., Hsieh, Colas, & Kanwisher, 2011; E. Yang, Blake, & McDonald, 2010), and thus compared to fixed duration of a stimulus under CFS, the breaking-CFS provides more flexible suppression (and also unconscious processing) time for individual difference.

Cons. (1) Since the breaking-CFS is a direct measure, it is difficult to distinguish the contribution from low-level and high-level information if a stimulus contains different properties at different levels (see above the discussion of fearful face that contains large eye-white regions, for example). (2) This measurement hardly uses objective criteria of unconscious processing, and thus begs similar criticisms for subjective criteria. That is, response bias may be one confounding factor the blurs the boundary between conscious and unconscious processing (Stein & Sterzer, 2014).

Physiology-CFS. The physiology-CFS is a direct measure and based on subjective and/or objective criteria. This paradigm measures some physiological indices, such as blood-oxygen-level dependent (BOLD) signal, event-related potentials (ERPs), and eye-movements, when stimulus is presented and suppressed under CFS. At the same time, subjective visibility and/or objective performance of the suppressed stimuli can be measured to support the claim that the stimuli are subliminal.

Pros. (1) Besides subjective criteria, this method can also adopt objective criteria to define unconscious processing. (2) Depending on the experimental procedure, it can use either fixed duration or flexible duration for the stimulus under CFS. (3) Physiological index provides additional information about how biological system processes different stimuli unconsciously.

Cons. (1) Physiology-CFS is a direct measure like breaking-CFS, it is thus also hard to distinguish the levels of processing if a stimulus contains different properties. (2) Technically speaking, dynamic masks in CFS would induce strong noise signals and thus induce low signal-to-noise ratio to measure physiological response in some circumstance.

Aftereffect-CFS. The aftereffect-CFS is an indirect measure and based on subjective and/or objective criteria for unconscious processing. This paradigm presents an invisible probe during CFS, and measures its influence on the subsequently presented target. Generally speaking, we define the "aftereffect" here that includes perceptual aftereffect (e.g. Gibson & Radner, 1937), semantic priming (e.g., Marcel, 1983) or attentional cueing (e.g., Posner, 1980). Similar as the physiology-CFS, this measure can also adopt both subjective and objective criteria to define unconscious processing.

Pros. (1) The criteria would be stringent for defining unconscious processing if both subjective and objective criteria are used. (2) By manipulating the congruency between probe and target at a specific level, it can clearly attribute to specific processing without low-level explanation (see the above discussion about numerical-priming effect for example: only numerosity processing can reveal the numerical distance effect),

Cons. (1) Selecting corresponding target for the invisible probe may be difficult in some situations, especially for high-level information. (2) Using a fixed duration to present an invisible probe would miss individual differences, and this shortcoming may induce unsuccessful suppression. (3) Different information may require different time to process under CFS, it thus hard to find a suitable duration for different stimuli (note: to address this issue, we therefore proposed the "contingent-CFS paradigm" in the "future direction" section, p. 91).

Goal and Overall View of this Thesis

As mentioned in the previous section, different measures have their own advantages and disadvantages, which also reflecting some criticizes from previous studies (Gray et al., 2013; E. Yang, Hong, et al., 2010). For example, using facial expression alone without finding different results between formal condition and control condition (e.g., Gray et al., 2013; E. Yang et al., 2007) that would be hard to differentiate whether unconscious processing of facial expression is due to emotional content or low-level properties. Moreover, since breaking-CFS is a subjective measure of unconscious processing, it is also arguable whether using breaking-CFS can provide unequivocal evidence that the processing is fully unconscious (T. Stein & Sterzer, 2014). Regarding the aftereffect-CFS, whereas some found facial-expression aftereffect (Adams et al., 2010), others argue that this result may involve partial consciousness (Yang, Hong, et al., 2010).

Motivated by these studies, in the current thesis we conducted three studies with different approaches to test whether emotional content of facial expression can be processed under CFS. In Study 1 and Study 2, we used the breaking-CFS paradigm to address this issue, and instead of using face alone, we manipulated the meaningfulness (e.g., familiarity and multisensory integration) of facial expression. Since meaningfulness may influence emotional content, but not low-level properties, we are thus more confident that the unconscious processing was specific to emotional content of facial expression. Furthermore, we adopted Physiology-CFS measurement (i.e., eye-movement recording) in Study 2 to exclude the possible explanation of unclear boundary between conscious and unconscious processing under breaking-CFS. In Study 3, we used priming-CFS paradigm and used both subjective and objective measurement to define unconscious processing. More specifically, we reciprocally used face and word as prime and target to exclude any low-level explanation in different experiments, and measured affective priming effect to further explore the possibility that emotional content can be processed under CFS.

Study 1. Since familiarity has been considered to affect emotional judgment of facial expressions but not low-level properties, we adopted different types of familiarity to test the modulation effect on unconscious processing of facial expression in this study. By measuring breaking-CFS, in Experiment 1, we adopted own-race familiarity (own-race face versus other-race face) on facial expression to manipulate the meaningfulness of emotional content. The impact of own-race familiarity on facial expression was also compared with the effect of orientation familiarity of faces (upright vs upside down; Experiment 2), which was used as a control condition in previous studies (e.g., T. Stein & Sterzer, 2012; E. Yang et al., 2007).

Study 2. This study adopted multi-sensory integration of emotional information to test unconscious processing of facial expression. We assumed that only if the emotional content of facial expression can be extracted under CFS, then it can be integrated with congruent affective voice. By measuring breaking-CFS, in experiment 1, we first tested whether face-voice can be integrated unconsciously without confounded low-level explanations. In Experiment 2, we tested whether face-voice can be integrated unconsciously based on specific emotion. Moreover, eye-movements were recorded to compare with the suppression time data.

Study 3. To better illuminate the emotional information can be extracted under CFS, we tested whether emotional meaning can interact with different types of stimuli as revealed by affective-priming effect. In Experiment 1, an invisible prime, either facial expression or emotional word, was presented for 1000ms under CFS, and an emotionally congruent or incongruent target was presented subsequently. In Experiment 2, we presented invisible prime for 200ms to test whether the priming effect would be different for a short duration. In Experiment 3, we changed the type of target to exclude the possibility of response bias.

Study 1

Own-Race Face, but not Face Orientation, Modulates Unconscious Processing of Emotional Expression under Interocular Suppression

(Submitted)

Abstract

While some studies suggest that *emotional content* of facial expression gains prioritized access to visual awareness, other studies suggest instead that can be entirely explained by *low-level* properties. Since familiarity has been considered to affect meaningfulness of emotional content but not low-level properties, this factor provides an opportunity to test unambiguously whether emotional content can be processed unconsciously. By adopting the continuous flash suppression (CFS) paradigm to render faces invisible, we manipulated the familiarity of face race and face orientation and measured the time for different facial expressions to get access to awareness. The results showed that although both race and orientation familiarity can influence the unconscious processing of *face*, only race familiarity can influence the unconscious processing of *facial expression*. This suggests that even when facial expressions are not consciously perceived under interocular suppression, their emotional content, which is modulated by race familiarity, can still be processed without awareness.

Introduction

In the recent decade, the continuous flash suppression (CFS; Fang & He, 2005; Tsuchiya & Koch, 2005) paradigm was largely used to explore how efficiently invisible information can gain access to awareness. In this paradigm, a series of dynamic high-contrast masks are presented to one eye, which renders a critical stimulus to the other eye invisible due to interocular suppression. By manipulating the contrasts between critical stimuli and masks, eventually the suppression breaks where the critical stimulus becomes visible again. The time required for breaking CFS (named b-CFS hereafter) is usually considered as an index of how efficiently a stimulus is unconsciously processed (Jiang et al., 2007). By adopting this method, several studies demonstrated that both low-level, such as spatial frequency (E. Yang & Blake, 2012), motion direction (Kaunitz, Fracasso, Lingnau, & Melcher, 2013), perceptual grouping (Wang, Weng, & He, 2012) and high-level information, such as familiarity (Gobbini et al., 2013; Jiang et al., 2007) and semantic meaning (Costello, Jiang, Baartman, McGlennen, & He, 2009; Y. H. Yang & Yeh, 2011) can be processed unconsciously.

Among the studies probing high level information, some have shown that familiar stimuli can be processed more efficiently than unfamiliar ones as indexed by faster b-CFS times. For example, Jiang et al. (2007) showed that upright faces were detected more quickly than inverted faces, which suggested that familiar orientation of faces has privileged detection under interocular suppression (see also Gray et al., 2013; E. Yang et al., 2007). In addition to faces, this inversion effect also extends to other stimuli such as words (Y.-H. Yang & Yeh, 2014; Y. H. Yang & Yeh, 2011) and bodies (T. Stein, Sterzer, & Peelen, 2012). Other than the impact of familiar orientation, previous studies also showed that familiar faces took less time to break interocular suppression; as demonstrated with the comparison between own-race and other-race faces (T. Stein, End,

& Sterzer, 2014) and between personally familiar and stranger faces (Gobbini et al., 2013). Taken together, these evidences suggest a preferential access to awareness of familiar information.

However, unlike the overwhelming evidence that familiar objects can be processed more efficiently under CFS, research on unconscious processing of emotional content in facial expression is still debated (Gray et al., 2013; T. Stein & Sterzer, 2012). This controversy is based on the observation that low-level properties (e.g., eye-white region) always contribute to representing facial expressions (e.g., fearful faces almost always contain larger eye-white region). It is thus difficult to dissociate the contribution of low-level features from emotional content when investigating the preferential access of different facial expressions under CFS. For example, E. Yang et al. (2007) found that fearful faces broke through interocular suppression more quickly than neutral and happy faces, but similar patterns were seen when faces were inverted or even when only the eye-regions of different facial expressions were presented. Although the authors concluded that salient eye-white regions embedded in fearful faces contributed to the preferential access of fearful to neutral/happy faces, it remains unclear whether emotional content and/or low-level properties (e.g. high contrast) of eye-white region contributed to that result.

In order to dissociate low-level and high-level properties of facial expressions, Gray et al. (2013) used inverted orientation combined with reversed polarity of luminance to generate a control face condition, which interrupted emotional information more severely than inversion alone but kept the low-level properties similar to the original face condition (upright and unaltered polarity of luminance). Their results showed that the control face condition revealed similar pattern as the original face condition; that is, fearful faces gained preferential access to awareness compared to neutral and happy faces. Therefore they concluded that given the higher strength of emotional content disruption in their control, the preferential access for fearful faces under interocular suppression could be entirely explained by the *low-level properties* embedded in emotional expression.

While contributions by low-level properties seem sufficient to explain preferential access of facial expression, the possibility that emotional content can undergo unconscious processing is still not completely eliminated. Note that in the first experiment from Gray et al. (2013), on top of the finding that the accuracy of explicit emotional judgment of inverted faces (with unaltered polarity of luminance) was higher than chance level, inverted faces with emotional content also gained privileged access under CFS. As suggested by Gray et al. (2013), the similar results between upright and inverted faces could be due to low-level properties (e.g., high contrast of eye-white region in fearful face) which were not affected by inversion; or, alternatively, emotional content per se could be extracted even for inverted faces. In other words, manipulating the face orientation alone is insufficient to serve as an unambiguous control to dissociate low-level and high-level information of facial expressions.

To distinguish the contribution of emotional content from low-level properties in the processing of facial expression, one alternative way is to pinpoint certain factors that only affect the processing of emotional content without changing low-level properties (c.f. Sterzer, Hilgenfeldt, Freudenberg, Bermpohl, & Adli, 2011). In the literature of emotional processing, race familiarity has been shown to modulate the emotional recognition of facial expression (Elfenbein & Ambady, 2002; Wild-Wall, Dimigen, & Sommer, 2008). For example, although the recognition of emotional expression is thought to be inborn (Ekman & Friesen, 1971; Izard, 1994), the effectiveness of presenting and perceiving emotional expression is greatly dependent on the subjects' native culture (Matsumoto, 1989). One previous study also revealed that emotion recognition of facial expression is better in faces from the same race as the participants (own-race faces) than in faces from other races (other-race faces) (Elfenbein & Ambady, 2003). Since race familiarity should not influence the processing of low-level properties (e.g., contrast) yet can influence emotion recognition, it is potentially a reliable factor to test whether emotional content can be processed unconsciously under CFS.

The main purpose of the present study is to investigate unconscious emotion processing by introducing race familiarity as a variable to modulate facial expression under the CFS paradigm. We will conduct two sets of experiments in this study. In the first experiment, we will test whether race familiarity (own-race faces versus other-race faces) is capable of modulating unconscious processing of facial expression. If differences in b-CFS times between disparate facial expressions (fearful, happy, neutral) were found to be more discrepant for own-race than for other-race conditions, we could infer that emotional content of facial expression can be processed unconsciously since low-level properties would be relatively constant between own-race face and other-race face conditions. In the second experiment, we will attempt to replicate the result from previous studies (e.g., E. Yang et al., 2007) that adopted the face inversion paradigm (upright face vs inverted face) so as to contrast effects of different types of familiarity with the result from our first experiment. We predict that unconscious processing of face can be affected by familiarity (race familiarity and familiar orientation) in general, but the pattern of differential *emotional expression* processing would be identical across the two different orientations.

Experiment 1

Previous study has shown that Caucasian participants detected Caucasian faces (own-race) more quickly than black faces (other-race) under CFS (T. Stein et al., 2014), which suggested that visual awareness of face can be shaped by the experience of social contact. In this experiment, we tested whether this own-race effect can modulate processing efficiency of different emotional expressions (happy, neutral, and fearful). To generate own-race effect, we recruited Taiwanese students as participants and used the FaceGen software (Singular Inversions Inc.) to create Asian faces (own-race) and Caucasian faces (other-race).

The predictions of this experiment are as follows. First, if the previous own-race effect observed with Caucasian participants can be generalized to other races, Asian faces should also be detected faster than Caucasian faces in Taiwanese participants. Second, if own-race familiarity can modulate the unconscious processing of emotional expressions, the b-CFS time differences between differential emotional expressions should be larger in own-race faces than other-race faces.

Method

Participants. Twenty Taiwanese participants were recruited in this experiment. All had normal or corrected-to-normal vision. Ethical approval for this study was granted by the Ethics Committee of the Department of Psychology, National Taiwan University.

Stimulus materials and apparatus. Visual stimuli were presented on a 21-inch Eizo T966 CRT monitor (1024 × 768 resolution at 60Hz refresh rate) and controlled by Matlab 2012a (the Math Works, Natick, USA) with Psychophysics Toolbox 3.0 (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007) on a Microsoft Windows PC. Participants viewed stimuli through a four-mirror stereoscope, such that two fusion contours (10.623° × 10.63°, 0.3° thickness with white noise pattern) with gray background (RGB values of 127, 127, 127) could be respectively presented to different eyes to support stable fusion. A white fixation-cross (0.66° × 0.66°; RGB values of 255, 255, 255) was displayed in the center of each fusion contour, and participants were asked to focus on the fixation throughout every trial. A face (4.5° × 4.5°) was presented on one of the four quadrants in one fusion contour (center-to-center distance: 3.3°), and the Mondrian-like masks that contained 1000 small patches (random size: 0.02° to 1.07°; random gray scale: 0~255) were presented in the other fusion contour. To decrease response variance between dominant eye and non-dominant eye, the eyes presented with the face or the masks were kept constant within subject, and were counterbalanced between subjects.

To generate Asian faces and Caucasian faces with similar low-level features, we adopted the FaceGen software (Singular Inversions Inc.) to generate virtual faces with 16 posers. Eight Asian (4 females) and eight Caucasian (4 females) posers were generated, and each poser had three facial expressions (happy, neutral, and fearful). To avoid abrupt onset triggering a dominant face percept at the beginning of the trial, the contrast of the face was ramped up from 0 % to 100 % within the first second (E. Yang et al., 2007). The contrast of Mondrian-like masks stayed at 100% for the first second and was ramped down from 100% to 0% within the remaining 5 seconds.

Rating task was conducted following the b-CFS task to assess the subjective evaluation about the race and the valence of manipulating faces. In the rating task, the face stimuli (4.5° × 4.5°) was presented in the upper visual field of the center display (center-to-center distance: 5°). A number line (0.5° × 19°) was presented in the vertical center of display. To represent the rating scale, seven small bars (1° × 0.5°) were presented on the number line with equal distance. Another white square (1° × 1°) was used as the index of rating position, and its position could be adjusted by the left- and right-arrow keys.

Design. This experiment contained two parts, one was the b-CFS task, and the other was the rating task. In the b-CFS task, a within-subjects design with race (Asian and Caucasian) and facial expression (happy, neutral, and fearful) was conducted. Face race, facial expression, visual quadrant for face presentation, and face poser were orthogonally manipulated and repeated twice, adding up to 384 trials in total. The order of above combined condition was pseudo randomized to keep an equal number of trials in each condition.

In the rating task, the 48 faces (16 posers, each presented 3 facial expressions) were randomly presented. During the face presentation, a race-rating question and a valence-rating question were introduced sequentially.

Procedure. In the b-CFS task, participants pressed the 'z' key to start each trial. A face was presented to one selected eye, while Mondrian-like masks flashing at 10 Hz were presented to the other eye. Participants were instructed to press the 'z' key again as soon as possible if they detected any part of a face. If they did not press the key, the trial would stop automatically after 6 seconds. After that, they were also required to judge the side of the screen where the face was presented by pressing the 'o' or the 'p' key for left or right visual fields, respectively (See Figure 1.1).

After the b-CFS task, the participants were instructed to rate the race and the valence of the faces with a 7-points Likert scale. They pressed the left and the right arrow keys to rate the race and the valence from 1 to 7 without time limit. In the race-rating task, number 1 referred to Caucasian, number 4 referred to ambiguous race, and number 7 referred to Asian on the scale. In the valence-rating task, number 1 referred to negative valence (fearful), number 4 referred to ambiguous valence, and number 7 referred to positive valence (happy) on the scale.

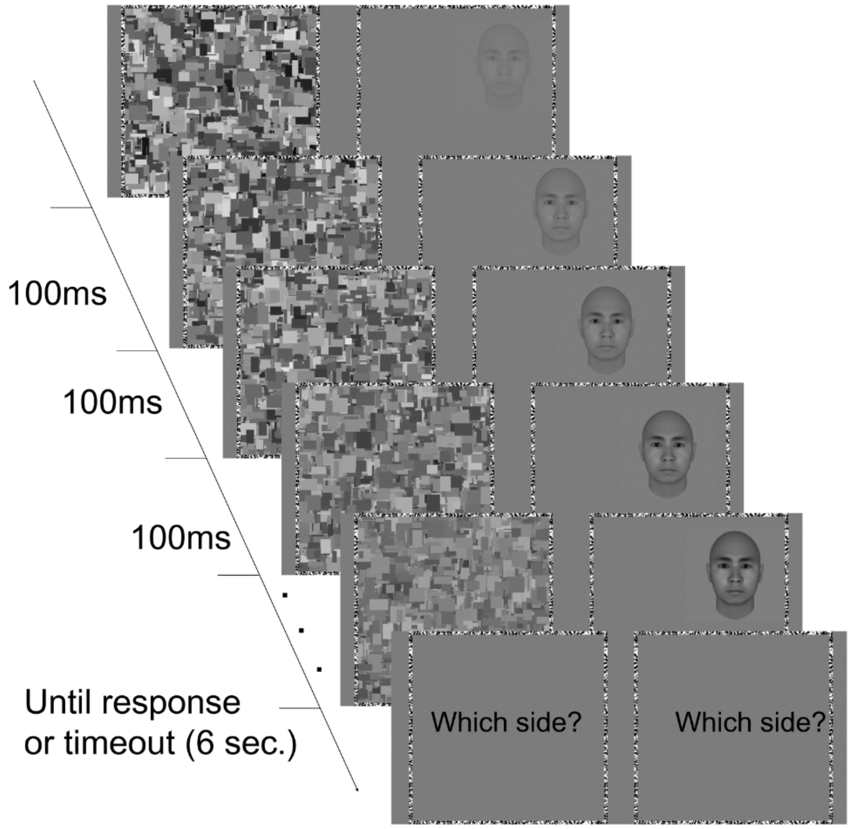

Figure 1.1. Stimuli and procedure. The face and the masks were presented dichoptically for 100 ms per frame. The contrast of faces increased gradually from 0 % to 100 % within the first second and kept constant until the end of that trial. The contrast of masks was ramped down from 100% at 1s after trial onset to 0% over 5s. The trial stopped when participants pressed a key to indicate detection of the face, or after 6s if no response was made. After detection, participants answered which side they perceived the face.

Results

B-CFS task. In the b-CFS task, the suppression time of trial with error response of location judgment were excluded from the analysis (*Mean accuracy* = 99.01 %, *SD* = .09 %). A two-way repeated measures analysis of variance (ANOVA) with the factors race (Asian and Caucasian) and facial expression (happy, neutral, and fearful) showed a significant main effect of race [$F(1,19) = 31.202$, $MSE = .008$, $p < .001$, $\eta_p^2 = .622$], reflecting shorter suppression time for Asian faces (*M* = 1.433, *SD* = .231) than Caucasian faces (*M* = 1.527, *SD* = .276), and a significant race-by-facial expression interaction [$F(2,38) = 3.731$, $MSE = .002$, $p = .033$, $\eta_p^2 = .164$]. There was no significant main effect of facial expression [$F(2,38) = 1.278$, $p = .299$, $\eta_p^2 = .063$]. The simple main effect of facial expression in the Asian face [$F(2, 76) = 4.324$, $MSE = .003$, $p = .0166$, $\eta_p^2 = .102$] revealed that happy face (*M* = 1.407, *SD* = .218) was detected faster than fearful face (*M* = 1.454, *SD* = .234) by Tukey test. There was no significant simple main effect of facial expression in the Caucasian face [$F(2, 76) = .511$, $p = .602$, $\eta_p^2 = .0146$] (See Figure 1.2).

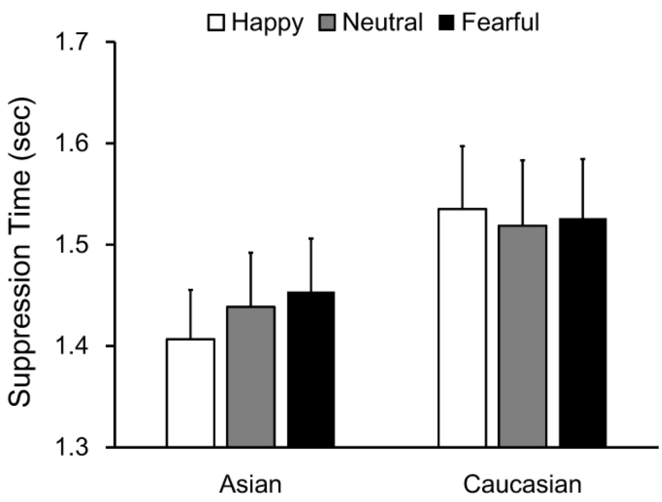

Figure 1.2. Mean suppression time in Experiment 1.

Rating task. In the rating question of "which race?", a two-way repeated measures ANOVA with the factors of race and facial expression revealed a significant main effect of race [$F(1,19) = 115.931$, $MSE = 2.678$, $p < .0001$, $\eta_p^2 = .859$], the Asian face (*M* = 5.531, *SD* = .931) was rated more "Asian" than Caucasian face (*M* = 2.315, *SD* = .776). The main effect of valence [$F(2, 38) = 2.028$, $p = .146$, $\eta_p^2 = .096$] and interaction [$F(2, 38) = 2.260$, $p = .118$, $\eta_p^2 = .106$] were not significant (See Figure 1.3a).

In the rating question of “which valence?” a two-way repeated measures ANOVA with the factor race and facial expression revealed a significant main effect of facial expression [$F(2,38) = 304.644$, $MSE = .373$, $p < .0001$, $\eta_p^2 = .941$]. The Tukey test showed that happy face ($M = 5.719$, $SD = .481$) was rated more positively than neutral face (M = 3.994, SD = .370), which was rated more positively than fearful face ($M = 2.35$, $SD = .530$). The main effect of race was also significant [$F(1,19) = 14.685$, $MSE = .043$, $p = .001$, $\eta_p^2 = .436$], the Asian face ($M = 4.094$, $SD = 1.465$) was rated more positively than the Caucasian face ($M = 3.948$, $SD = 1.435$). However, there was no interaction between race and facial expression [$F(2,38) = .721$, $p = .493$, $\eta_p^2 = .037$] (See Figure 1.3b).

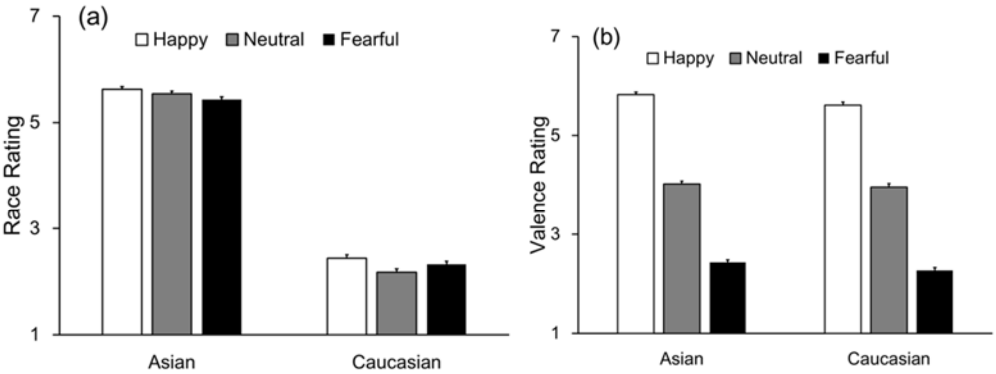

Figure 1.3. Rating results in Experiment 1. Result of race rating task: scale 1 referred to Caucasian face and scale 7 referred to Asian face (Figure 1.3a). Result of valence rating task: scale 1 referred to negative valence and scale 7 referred to positive valence (Figure 1.3b).

Correlation analysis. To further reveal the suppression time as the function of race and facial expression, we calculated the correlation coefficient between suppression time and valence rating in different races of face. We used scale 4 in race rating question as cutting point to sort the 48 faces (16 posers, each presented 3 facial expressions) into Asian faces and Caucasian faces. The result showed that participants can clearly distinguish 24 Asian faces (M = 5.55, SD = .24) and 24 Caucasian faces (M=2.37, SD =.57) from our manipulation [$t(23) = 20.65$, $SE = .154$, $p < .001$]. Based on this sorting, we found that the suppression time was significantly correlated with the facial expression in Asian faces (*Pearson's r* = -.473, $p = .02$) but not in Caucasian faces (*Pearson's r* = .024, $p = .913$). This result suggests close relationship between the suppression time and the facial expression in own-race faces but not in other-race faces (See Figure 1.4).

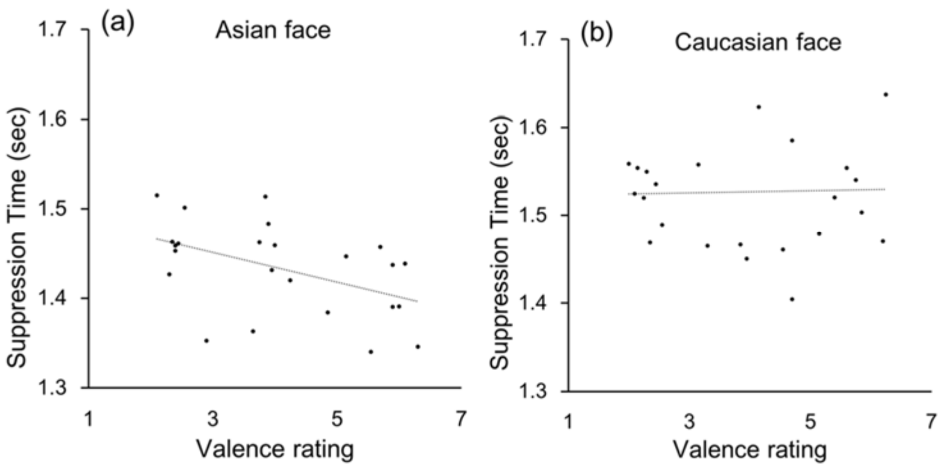

Figure 1.4. Correlation between suppression time and valence rating (1 = negative valence and 7 = positive valence) in Experiment 1. Each dot represented one face image. Suppression time was negatively correlated with positive valence in Asian face (Figure 1.4a). There was no significant correlation between suppression time and valence in Caucasian face (Figure 1.4b).

Discussion

The result of this experiment showed that Taiwanese participants detected Asian face more quickly than Caucasian face, which replicated previous finding of own-race effect with Caucasian participants (T. Stein et al., 2014). On top of this observation, we also found that processing of emotional expression is modulated by race familiarity; that is, the suppression times of emotional expressions were different only when viewing

Asian face and not when viewing Caucasian face. Parallel to this finding, the correlation between suppression times and valence rating was only significant in Asian face and not in Caucasian face. This convergent evidence indicates that unconscious processing of emotional expression, which was defined by our manipulation and subjective rating, was modulated by the own-race familiarity. In line with our hypothesis that emotional content, but not low-level properties, should be affected by the own-race effect, this result also implies that the former of facial expression can be extracted under interocular suppression state.

Regarding the modulation effect of familiarity on the unconscious processing of emotional expression, previous studies also adopted inverted face as control to test whether unfamiliar orientation of face can trigger different pattern compared to the upright face (Gray et al., 2013). However, most studies found that inverted face revealed similar pattern as upright face, and this effect was mainly due to high saliency of eye-white region in fearful face. Since we found that happy face was detected faster than fearful face in the current study, the strong eye-white region embedded in fearful face should not be the only factor to determine the processing efficiency in b-CFS. This observation provided an alternative possibility that inverted face would reveal a different pattern to upright face if local features in fearful face were not as critical in as previous studies. Hence in the next experiment, we investigated the effect of the familiarity of face orientation using face inversion to test this possibility.

Experiment 2

In this experiment, we tested whether face inversion can also modulate the unconscious processing of facial expression under interocular suppression. If different b-CFS times between disparate facial expressions were found to be weaker in inverted face than upright face, it would suggest that face inversion can also serve as a control to interrupt the emotional content of facial expression (see Gayet, Van der Stigchel, & Paffen, 2014 for the discussion of control condition under b-CFS). On the other hand, if similar patterns were found in both upright and inverted condition, we may deduce that face inversion is not a suitable baseline control condition, as suggested by Gray et al. (2013).

In addition, the result from the first experiment showed that happy face gained preferential detection, which contradicted previous findings that supported preferential detection for fearful face (e.g., E. Yang et al., 2007). To better support the result of the first experiment, we adopted FaceGen (Singular Inversions Inc; as in Experiment 1) and TFEID (Taiwanese facial expression image database; L. F. Chen & Yen, 2007) to represent facial expressions in Experiment 2a and 2b, respectively. Compared to the FaceGen which is a database of virtual faces, TFEID is a database that includes real faces from Taiwanese posers, and the reliability of different facial expression has been supported by group-rating. By incorporating this naturel face database, which could increase ecological validity, we test whether similar result as the first experiment for valences can also be found.

Method

Participants. According to the previous study (E. Yang et al., 2007), we recruited a similar number of participants (6~12 observers among the three experiments in their study) in this experiment: nine in Experiment 2a and nine in Experiment 2b.

Stimulus materials and apparatus. The apparatus and stimulus display were the same as in Experiment 1 except for the following. In addition to the face stimuli generated by the FaceGen software (Singular Inversions Inc.) in experiment 2a, we also adopted faces from the TFEID (L. F. Chen & Yen, 2007) in experiment 2b. Both included eight posers (four females), each posing three facial expressions (happy, neutral, and fearful). The stimuli were controlled by E-Prime software (Psychological Software Tools, Pittsburgh, PA).

Design. A within-subjects experiment with face orientation (upright and inverted) and facial expression (happy, neutral, and fearful) was conducted. Face orientation, facial expression, visual quadrant for face presentation, and face poser were orthogonally manipulated and repeated twice, yielding 384 trials in total. The order of the above combined condition was pseudo randomized.

Procedure. The procedure of the b-CFS experiment was exactly the same as the first experiment but without the rating tasks.

Results

The suppression time with correct judgment of face location were submitted for further analysis. The accuracy of location judgment in experiment 2a and 2b was 98.76 % (*SD* = 1.24 %) and 98.78 % (*SD* = .82 %), respectively. A two-way repeated measures ANOVA with the factor of face orientation (upright and inverted) and facial expressions (happy, neutral, and fearful) was conducted.

Experiment 2a. The main effect of face orientation [$F(1,8) = 37.121$, $MSE = .016$, $p < .001$, $\eta_p^2 = .823$] and facial expression [$F(2,16) = 17.869$, $MSE = .008$, $p < .001$, $\eta_p^2 = .691$] were significant. Upright face ($M = 1.642$, $SD = .450$) was generally detected more quickly than inverted face ($M = 1.850$, $SD = .481$). The Tukey test showed that happy face ($M = 1.653$, $SD = .449$) was detected more quickly than neutral face ($M = 1.751$, $SD = .481$), which was detected faster than fearful face ($M = 1.833$, $SD = .485$). There was no significant interaction between face orientation and facial expression [$F(2,16) = .202$, $p = .819$, $\eta_p^2 = .025$] (see Figure 1.5a).

Experiment 2b. The main effect of face orientation [$F(1,8) = 26.155$, $MSE = .091$, $p < .001$, $\eta_p^2 = .766$] and facial expression [$F(2,16) = 13.719$, $MSE = .033$, $p < .001$, $\eta_p^2 = .632$] were significant. Upright face ($M = 1.787$, $SD = .525$) was generally detected more quickly than inverted face ($M = 2.207$, $SD = .707$). The Tukey test showed that both happy face ($M = 1.842$, $SD = .526$) and neutral face ($M = 1.988$, $SD = .684$) were detected more quickly than fearful face ($M = 2.161$, $SD = .709$). There was no significant interaction between face orientation and facial expression [$F(2,16) = 2.754$, $p = .094$, $\eta_p^2 = .256$] (see Figure 1.5b).

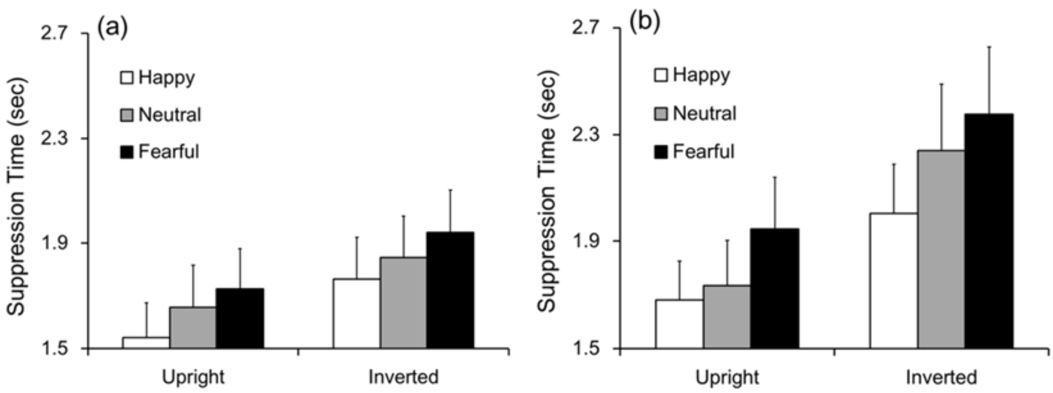

Figure 1.5. Mean suppression time in Experiment 2a (Figure 1.5a) and Experiment 2b (Figure 1.5b).

Discussion

The results from Experiment 2a and 2b consistently showed that upright face was detected more quickly than inverted face, which suggests that familiar orientation gains preferential access to awareness. In addition, happy face was detected more quickly than neutral and fearful face, which also replicated the result of the first experiment. This result implies that the local saliency embedded in fearful face is not the only factor to determine the unconscious processing of facial expression, since our result across different experiments showed that happy face gained preferential access to awareness under interocular suppression.

Compared to the interactive impact of own-race effect on the unconscious processing of facial expression, orientation familiarity did not have the same modulation effect as that of race familiarity. Due to the similar results in both upright and inverted faces in Experiment 2a and 2b, it is unclear how much contributions were made by each factor, the emotional content or the low-level properties. This observation implies that orientation familiarity seems unable to provide an unequivocal baseline to dissociate low-level and high-level information in the processing of facial expression.

General Discussion

From Experiment 1 and 2, our results showed that own-race face and upright face were detected faster than other-race face and inverted face, respectively. This result is consistent with previous findings that familiar information is processed more efficiently than unfamiliar information under CFS (Gobbini et al., 2013; Jiang et al., 2007; T. Stein et al., 2014). Moreover, both experiments clearly showed that happy face was detected faster than neutral and fearful face. Since the processing of facial expression was modulated by the own-race familiarity, this implies that emotional content of facial expression can be processed unconsciously under CFS. On the other hand, face inversion did not affect the unconscious processing of facial expression. These results demonstrated that although both familiarity (own-race familiarity and familiar orientation) can affect the access efficiency of face, only own-race familiarity can be a modulator to unconscious processing of emotional content of facial expression.

The finding that familiar identity can influence processing of facial expression has been observed in several studies. For example, Baudouin, Sansone, and Tiberghien (2000) found that recognition of facial expression was more accurate in famous face and familiar face than their counterparts. Moreover, Elfenbein and Ambady (2003) also found that accuracy of facial expression judgment favored own-race face than other-race face, and the other-race disadvantage on facial expression judgment can be improved by cultural exposure to the other-race group. To encode facial identity and facial expression, previous studies have suggested distinct visual pathways such as lateral fusiform face area (lFFA) versus superior temporal sulcus (STS) to represent the invariant (for facial identity) and variant (for facial expression) visual properties (Haxby, Hoffman, & Gobbini, 2000). However, a modified neural model for recognition of familiar faces also suggested that STS is relevant to dynamic features for face recognition (Gobbini & Haxby, 2007). Since STS is sensitive to variations in facial expression under interocular suppression (Jiang & He, 2006), it is interesting to know whether the modulation effect of own-race familiarity on unconscious processing of facial expression observed here can also modulate the activation of STS under interocular suppression in future studies.

To demonstrate the processing of emotional information under interocular suppression, it is important to incorporate some control conditions that interrupted emotional content but kept low-level properties similar as manipulated stimuli (Gayet et al., 2014). However, when adopting figure inversion as control to disrupt holistic processing of words (Kao, Chen, & Chen, 2010) and faces (Kanwisher, Tong, & Nakayama, 1998), different impacts on the unconscious processing of emotional word and facial expression under interocular suppression are revealed. For example, a demonstration of the inverted orientation as control in the processing of emotional word under interocular suppression (Y. H. Yang & Yeh, 2011) showed that emotional meaning of word can only be accessed in the upright presentation but not in the inverted presentation. By contrast, previous studies (Gray et al., 2013; E. Yang et al., 2007) and the current finding both suggest that the manipulation of face orientation cannot clearly distinguish the levels of processing of facial expression. The inconsistent impacts of inversion between words and faces could be due to that local feature is sufficient to represent facial expression (Whalen et al., 2004; E. Yang et al., 2007), but local stroke cannot represent the emotional meaning of words. This implies that though the inversion effect has potential to interrupt the meaning of information and serve as a control condition, its effectiveness highly depends on the stimuli used.

From our two experiments, the results clearly showed that happy face was detected faster than fearful face. Although this contradicted the "fearful prioritization effect" of emotional expression in previous studies (Tsuchiya, Moradi, Felsen, Yamazaki, & Adolphs, 2009; E. Yang et al., 2007), it was not necessarily odd in the literature of emotional face processing. For example, Calvo and Lundqvist (2008) found that happy face was identified faster and more accurately than fearful face under different viewing durations (25ms ~ 500ms), and similar result was found across different databases (Biehl et al., 1997; Tottenham et al., 2009). In addition, Hare, Tottenham, Davidson, Glover, and Casey (2005) manipulated fearful, neutral, and happy face in a go/no-go task, and asked participants to respond to a specific target face (e.g., fearful face) and to ignore all non-target faces (e.g., neutral and happy face). They found that happy face was responded to more quickly than fearful face in the go condition, and fearful face registered lower false alarm rate than happy face in the no-go condition. This result implies that processing efficiency of positive or negative emotion depends on approaching (go-task) or avoiding (no-go task) response. In the current study, the b-CFS time is a much better analogy to the target response in the go-task, since every facial expression served as target and the b-CFS time was measured when detecting any parts of the face. If this idea is supported, a prediction for approaching response in the current study should favor that happy face has faster detection than fearful face under interocular suppression, and this is indeed what we found.

The absence of the fearful prioritization effect in the current study could be partially caused by relatively lower impact of eye-white region in fearful faces than previous studies (e.g., Gray et al., 2013; E. Yang et al., 2007). For example, from the perspective of physical anthropology, most of Asian faces adopted here have "epicanthic fold" of the upper eyelid, which may present relatively smaller eye-region compared to the Caucasian face. Therefore, the small eye-white region in Asian facial expression may also decrease its saliency in fearful face. In addition, Eastern and Western observers may use different decoding strategies in face processing (Blais, Jack, Scheepers, Fiset, & Caldara, 2008) and facial expression (Jack, Blais, Scheepers, Schyns, & Caldara, 2009) , which may also bias specific action unit of facial expression in the b-CFS procedure. Further studies are needed to examine the interactive mechanism of transmitting and decoding facial expression (Smith, Cottrell, Gosselin, & Schyns, 2005) across different races when probing unconscious processing of emotional information by adopting CFS.

In conclusion, we found that race familiarity, but not orientation familiarity, can modulate the processing of facial expression under CFS. This finding shows that emotional content can be extracted without visual awareness. Whereas the processing of facial identity (e.g., own-race face) and facial expression are considered separately (Bruce & Young, 1986; Haxby et al., 2000), the current study supports the alternative perspective that there is interactive processing of familiar identity and facial expression (Baudouin, Sansone, & Tiberghien, 2000; Elfenbein & Ambady, 2003). Compared to the manipulation of face inversion that cannot provide unequivocal baseline, the current study provides a new way to test whether emotional content can be processed unconsciously, especially under CFS from which the debate of emotional content versus low-level features originated.

Study 2

Multisensory Integration of Emotion: Unconscious Binding of Face and Voice

(Submitted)

Abstract

Emotional expression can be perceived through either visual or auditory modality, or both. People who suffer from prosopagnosia or affective blindsight (i.e., not being able to recognize or see a face) can nevertheless perceive face and voice together to detect emotion. However, it remains unknown whether unconscious binding of facial expression and affective voice can also be observed in neurologically intact participants. We examine this by recruiting normal participants and using the continuous flash suppression paradigm to render the face stimuli invisible. Experiment 1 showed that participants could unconsciously integrate upright face with human voice. Experiment 2 supported that this unconscious audio-visual integration can be specific to emotional valence and eye-movement results reflected subsequent response of face detection. These results demonstrate for the first time that neurologically intact participants can experience unconscious multisensory integration of emotion and have important implications for the debate on whether facial expression can be processed unconsciously in an interocular suppression state.

Introduction

Emotional expression, which is critical for social life, can be perceived through at least two modalities. Facial expressions—as revealed by the direction of eye gaze, the movements of eyebrows, or the shape of mouth—provide emotional cues visually (Bruce & Young, 1986; Haxby et al., 2000). Affective voices—as delivered by variations in amplitude, stress, intonation, or pitch—provide emotional cues through audition (Belin, Fecteau, & Bedard, 2004; Scherer, 1995).

Although facial expression and affective voice are initially processed in visual and auditory pathways respectively, they can be integrated into a unified percept that affects how emotion is perceived (Adolphs, 2002; Belin et al., 2004; Campanella & Belin, 2007). For example, de Gelder and Vroomen (2000) paired a sad or happy voice with a face that was selected from a morphed expression ranging between different percentages of sad and happy faces. The participants were required to judge whether the facial expression they saw in each trial was sad or happy. Their results showed that a sad voice shifted the judgment of the morphed face toward sadder expression, and a happy voice shifted it toward happier expression (See also Dolan, Morris, & de Gelder, 2001). These results imply that adding an emotional cue from one modality (e.g., affective voice) can make a difference in how we appraise the emotional information in another modality (e.g., facial expression).

Some hints from patient studies seem to imply that such multisensory emotional information can be integrated unconsciously. For example, when a person with prosopagnosia who could not recognize faces was asked to rate the strength of affective voices, the facial expression that could not be explicitly recognized still affected the emotional rating of the voice (de Gelder, Pourtois, Vroomen, & Bachoud-Levi, 2000). Similarly, the "unseen" face could also affect the rating of emotional voice by affective blindsight patients, who claimed to be blind due to striate cortex lesion but could still discriminate facial expressions accurately by guessing (de Gelder, Pourtois, & Weiskrantz, 2002). These case studies imply that multisensory information is integrated before conscious awareness.

However, there are several review papers that discussed residual visual experience (Fendrich, 2001; Overgaard, 2011) or functional recovery in patient studies (Mogensen, 2011; L. Weiskrantz, 2000). For example, Weiskrantz (2000) reported that the blindsight patient G.Y. improved his awareness and sensitivity after repeated testing by several groups (Weiskrantz, 2000, p. 286). This finding leaves open the question that whether patient studies really probe fully unconscious processing of emotional information, Moreover, as reported by de Gelder, Morris and Dolan (2005), the patient G.Y. suffered the lesion of primary visual cortex at the age of seven years old, and therefore may have experienced some reorganization and recovery of his brain. Indeed, Bridge, Thomas, Jbabdi, and Cowey (2008) found that Patient G.Y. has strong anatomical differences with neurological intact controls at the connectivity between right LGN to contralateral area V5/MT, and substantial connection between bilateral V5/MT (Bridge et al., 2008). Based on these observations, it would be difficult to generalize the results from patient studies and infer unconscious processing in neurologically intact participants. By revealing the unconscious binding of multisensory emotional information in neurological intact participants, new insights can be provided not only on how the neural system integrates information from two modalities, but also the extent or depth of processing that can take place without consciousness.

Hence, our primary goal here is to examine whether normal participants process the binding of facial expression and affective voice unconsciously. To present visual stimuli (i.e., faces) for a relatively long duration that can lead to its integration with voice but lack awareness of the face (and thus the relation between face and voice), we adopted the continuous flash suppression paradigm (CFS; Fang & He, 2005; Tsuchiya & Koch, 2005). The CFS paradigm refers to a method in which continuously flashing masks are presented to one eye that interocularly suppresses the critical visual information presented to the other eye so that the critical visual information is invisible but has nonetheless been processed. Compared to the other paradigms used to block information from consciousness (see Kim & Blake, 2005 for review), this paradigm renders the critical information completely invisible for a relatively longer period of time. It is an excellent tool to probe the unconscious processing of various kinds of information, including afterimage (Tsuchiya & Koch, 2005), tools (Fang & He, 2005), numbers (Bahrami et al., 2010), gaze (Y. C. Chen & Yeh, 2012; T. Stein, Senju, Peelen, & Sterzer, 2011), and words (Costello et al., 2009; Jiang et al., 2007; S. Y. Lin & Yeh, 2015; Y. H. Yang & Yeh, 2011). By adopting CFS, we presented a face stimulus that was interocularly suppressed and paired with congruent or incongruent voice to see whether multisensory integration of emotional information can be processed unconsciously.

Our second goal was to provide a solution to the debate on whether facial expression can be processed unconsciously in an interocular suppression state. It has been questioned that the claim of unconscious processing of facial expression under CFS may be due to the distinct low-level features embedded in emotional expressions, such as the strong local contrast or large eye whites in fearful faces (Gray et al., 2013; E. Yang et al., 2007). Indeed, several low-level factors (e.g., spatial frequency, contrast, and orientation) are known to influence the suppression strength in dichoptic viewing (Fahle, 1982a, 1982b; Mueller & Blake, 1989; E. Yang & Blake, 2012). The important question is how to determine whether these studies really probed the emotional content as they suggested. The covariance between the low-level features and facial expressions seems to be an inevitable issue when manipulating different facial expressions. For example, E. Yang et al. (2007) found that the time required for breaking CFS (named b-CFS hereafter) in fearful faces were faster than happy and neutral faces. However, the larger eye whites of fearful faces than neutral/happy faces might have contributed to their results, because fearful faces were also detected faster even when the faces were inverted. To test the idea that perhaps low-level features contributed to previous results, T. Stein and Sterzer (2012) used similar features (e.g., eyebrows, mouth, and eyes) but with different orientations to construct schematic emotional faces. They found that the parallel curvature between the mouth and the face contour of a happy face—i.e., the low-level features but not emotional information—contributed to the faster detection of happy faces.

To avoid this inevitable confounding of low-level features when different facial expressions are manipulated, we take a different approach: pairing face and voice and examining whether there exists a congruency effect of the face-voice pair when they carry the same emotional information. If the facial expression we manipulated is unconsciously processed up to the emotional level, it is expected that it will be integrated with the congruent affective voice and lead to faster b-CFS time than the incongruent pair. On the other hand, low-level features that do not lead to emotional processing should not have different results when they are paired with the assumed "congruent" or "incongruent" affective voice.

As multisensory integration can occur at different levels, in this study, we conducted two experiments to examine whether voices with congruent facial information (Experiment 1) or congruent emotional information (Experiment 2) facilitate the detection of an invisible face via multisensory enhancement. In Experiment 1, we first established the basic finding that human voice can integrate with face and also used various baselines to exclude low-level feature explanations. In Experiment 2, we further manipulated different affective voice with facial expression to examine unconscious multisensory integration of emotion.

Experiment 1

We first tested the possibility of audiovisual binding by examining whether face-voice integration can be processed unconsciously. Specifically, we aimed at ruling out the possibility of low-level auditory facilitatory effect on visual processing. It has been established that a pure tone can enhance the brightness of visual stimulus and facilitate visual processing (e.g., Y. C. Chen & Yeh, 2008; B. E. Stein, Nancy, Wilkinson, & Price, 1996). More relevantly, our previous study demonstrated that a beep-sound speeds up visual information that is paired arbitrarily using the CFS paradigm (Y.-H. Yang & Yeh, 2014). Therefore, it is necessary to tease apart the face-voice integration being probed here from low-level auditory facilitatory effects on visual processing. To probe this issue, we used human faces and paired them with human voices in this experiment.

We also established different baselines in visual and auditory modalities to ensure that all possible controls were adopted for a fair estimation of the visual-auditory integration. In vision, we manipulated the familiarity of facial orientation by adding inverted faces as a control to compare with upright faces. Previous studies have revealed that inverted faces interfere with configural processing of face, and interrupt face identification and emotional judgment (Rhodes, Brake, & Atkinson, 1993; Sato, Kochiyama, & Yoshikawa, 2011; Valentine, 1988). This suggested that the inverted face may serve as an ideal baseline to exclude the familiarity of face percept and yet reserve most physical properties. In audition, a beep sound was used as a baseline to compare with the human voice to exclude the possibility that any multisensory integration observed is due to a low-level auditory facilitatory effect on visual processing as mentioned earlier. If the effect of face-voice integration was due to low-level integration, then human voice and beep sound should get similar results. We also added a non-sound condition as the baseline of auditory stimuli, to exclude the possibility that the strength of beep sound is too weak to exert any effect if we find that it cannot facilitate the processing of upright face. By including these two auditory baselines (i.e., beep sound and non-sound), we were able to compare different facilitatory effects of human voice and beep sound under upright and inverted face conditions, respectively.

We presented voices with invisible faces under CFS, and measured the release-from-suppression time of the faces. By manipulating different baselines in vision and audition, our prediction is that if human voice has a unique role to facilitate the detection of invisible faces under CFS, it will lead to the shortest reaction time (RT) in the upright face condition compared to the two baseline conditions (i.e., beep and no-sound conditions).

Method

Participants. Thirty-nine undergraduate students of National Taiwan University (NTU) participated in this experiment for course credit. They had normal or corrected-to-normal vision and self-reported normal hearing, and were naïve about the purpose of this experiment. All gave informed consent before the experiment. The ethics committee at the Department of Psychology, NTU, approved the experiment.

Stimulus, materials, and apparatus. Stimuli were presented via E-Prime software (Psychological Software Tools, Pittsburgh, PA) with an IBM compatible personal computer and presented on a 22-inch ViewSonic P225f CRT monitor (1024×768 resolution at 60Hz refresh rate).

The display contained two frames with one at the left side and one at the right side. The frame on the left side was projected to the left eye and right side to the right eye through a four-mirror-stereoscope. Each frame was surrounded by an outer square with random dots (10.70° × 10.70°, 0.18° thickness) and the two outer squares were used for facilitating binocular fusion throughout the trial. In one frame, a face was presented on one of the four quadrants. The other frame was filled with masks that consisted of rectangles of random sizes (from 0.02° to 1.07°) at random gray levels (from 0 to 255 on RGB scale) and changed every 100ms (10Hz). The face was presented pseudo-randomly with an equal chance for either of the two eyes and also an equal chance for one of the four quadrants in the visual field to balance the participants' ocular dominance and the display position

The contrast of the face was ramped up from 0% to 100% in 1000 ms to prevent an abrupt onset that would trigger the face stimulus being seen. The contrast of the masks was ramped down from 100% to 10% in 4500 ms so that the participants could detect the face within a certain range of time. While the contrasts of face and masks reached the maximum and minimum contrast, respectively, they were kept constant until the end of the trial. To prevent the participants from having any rhythmic expectation, the presentation of the face also jittered from 0 to 400ms at the onset of each trial.

Two faces were generated by the FaceGen (Singular Inversions Inc.) software. Each face was extended 4.18° × 4.18° visual angle and presented 3.46° (center-to-center) from the center of the frame. The sounds (duration: 1000ms, 70dB) included human voice and beep sound (440Hz), which were synchronized with the onset of the presented face and presented by loudspeakers. The voices were adopted from freeSFX Sound Effect (http://www.freesfx.co.uk) with royalty free license.

Design. We manipulated facial orientation (upright and inverted) and sound type (human-voice, beep, and no-sound) in a 2 x 3 within-subjects factorial design. The inverted face with no-sound and beep-sound conditions served as different baselines.

Each of the six conditions contained 64 trials, and the 384 trials in total were mixed and conducted randomly.

Procedure. At the beginning of each trial, the participant was asked to fuse the two frames using a four-mirror-stereoscope. By pressing the "z" key to start each trial, the participants were asked to press the same key again as soon as they detected any part of the face from CFS, and they did not need to know what they perceived (Jiang et al., 2007). The trial stopped when the key was pressed after detection or it would stop automatically if no response was made within 6 sec. After the detection response, the participants were required to press the "o" or "k" key to judge the location of the face that was presented either at the top or bottom visual field, respectively (see Figure 2.1).

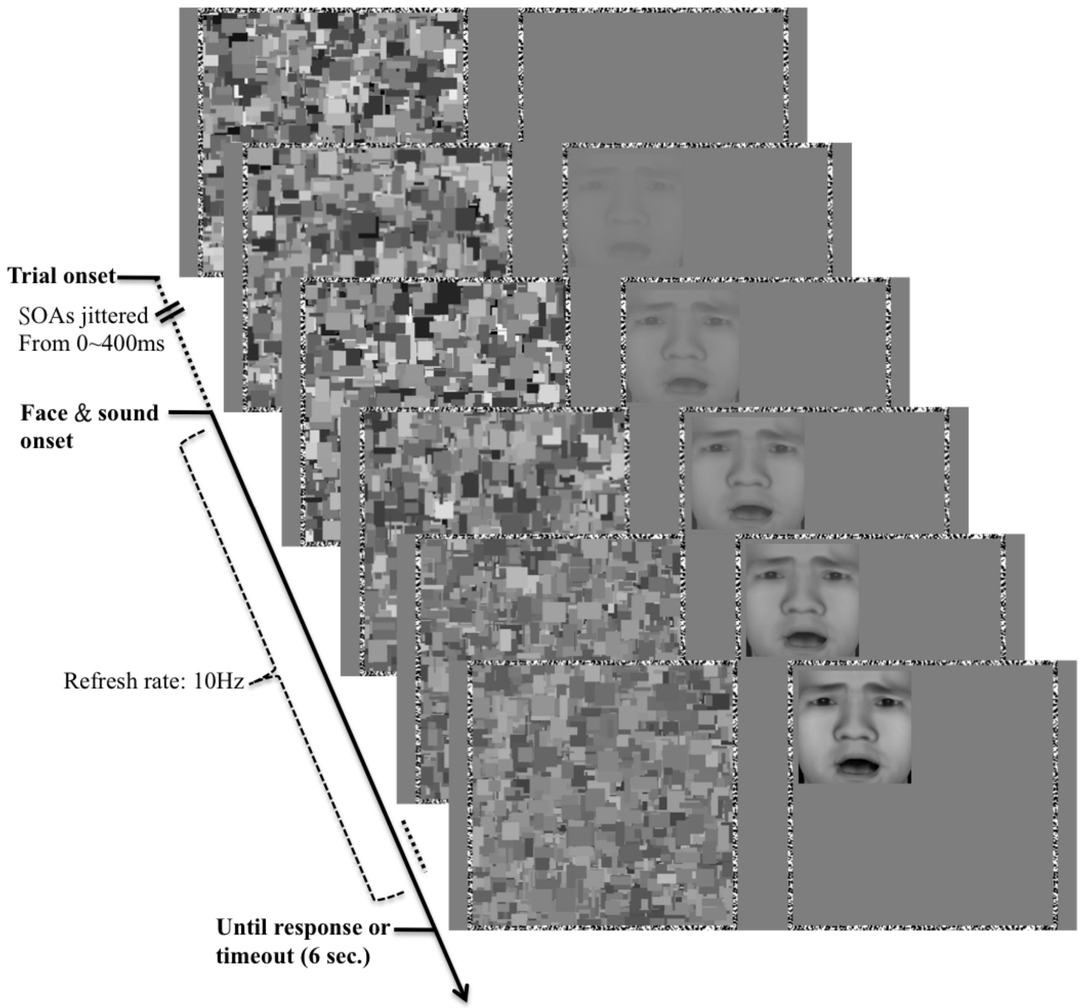

Figure 2.1. Stimuli and procedure of Experiment 1.

Results

Data from two participants were excluded due to their unsuccessful suppression of the continuously flashing high-contrast masks which made their RTs shorter than 500ms in more than half of the trials. This result suggests that there are some individual differences in successful suppression of CFS (cf., Hsieh et al., 2011; E. Yang, Blake, et al., 2010). RTs were calculated as the time of key press minus the onset time of the face. Only correct trials of the location judgment (mean accuracy = 97.58%, SD = 3.46%) were included for further analysis. Trials in which RTs were 3 standard deviations away from the sample mean were excluded. We conducted a 2 (facial orientation: upright and inverted) x 3 (sound type: human-voice, beep and no-sound) two-way repeated-measures analysis of variance (ANOVA) on the correct RTs.

The main effect of facial orientation was significant [$F(1,36) = 118.076$, $MSE = .063$, $p < .0001$; $\eta_p 2 = .766$]; upright faces ($M = 1.816$ s) were detected faster than inverted faces ($M = 2.183$ s). The main effect of sound was also significant [$F(2,72) = 12.562$, $MSE = .013$, $p < .0001$; $\eta_p 2 = .259$]; the no-sound condition ($M = 2.051$ s) had slower RTs than the voice or the beep condition ($M = 1.958$ s and $M = 1.988$ s, respectively; Tukey's test, *ps* < .01). The interaction between facial orientation and sound was marginally significant [$F(2,72) = 2.591$, $MSE = .011$, $p = .082$; $\eta_p 2 = .067$]. The sound types had different effects in the upright [$F(2,144) = 7.408$, $MSE = .012$, $p < .001$; $\eta_p 2 = .093$] and inverted conditions [$F(2,144) = 8.649$, $MSE = .012$, $p < .001$; $\eta_p 2 = .107$]. In the upright condition, the Tukey test showed that the voice condition ($M = 1.761$ s) had faster RTs than the beep ($M = 1.827$ s; $p < .05$) and no-sound ($M = 1.858$ s; $p < .01$) conditions. However, in the inverted-face condition, both the voice and beep conditions ($M = 2.155$ s; $M = 2.149$ s, respectively) had faster RTs compared to the no-sound condition ($M = 2.244$ s; *p*s < .01 in both conditions). (see Figure 2.2)

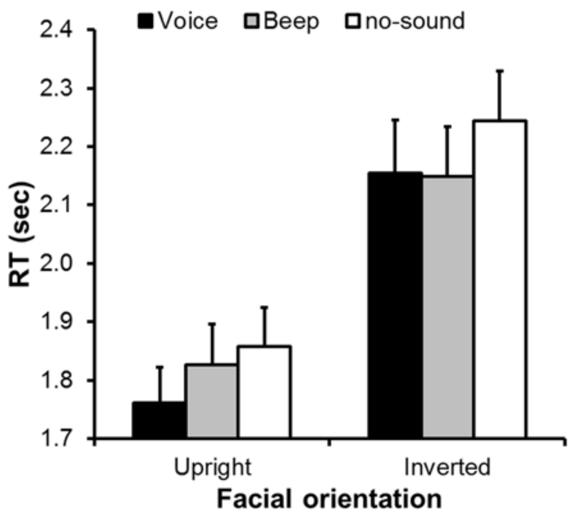

Figure 2.2. Results of Experiment 1. Upright faces were detected faster than inverted faces. The face in the no-sound condition was detected slower than that in the voice or beep conditions. Only voice facilitated the detection of an upright face, and both voice and beep facilitated the detection of an inverted face.

Discussion

We found that participants detected upright faces faster than inverted faces under CFS, indicating that upright faces are more familiar and meaningful for the participants thereby speeding up their unconscious processing. This result is consistent with previous studies (Jiang et al., 2007; Zhou, Zhang, Liu, Yang, & Qu, 2010), and implies different processing between upright and inverted faces. When different sound effects in the upright-face and inverted-face conditions were evaluated, we found that only the voice facilitated the detection of upright faces, but both voice and beep facilitated the detection of inverted faces compared to the no-sound condition. Since upright faces can only be integrated with human voice but not beep-sound under the CFS paradigm, we conclude that unconscious multisensory integration observed here is not due to low-level feature integration.

This finding is parallel to de Gelder, Vroomen, and Bertelson (1998)'s study, in which multisensory interaction of emotion could be found only with an upright face but not with an inverted face. Here we found that participants integrate the face-voice pairing unconsciously, implying multisensory integration of facial information is an automatic perceptual processing. We also found that the beep-sound condition was faster than non-sound condition with an inverted face. This result suggests that the null result of beep sound in the upright face condition is not due to the ineffictiveness of the beep sound. This is an important manipulation check to support that the beep-sound used here is efficient enough to represent the low-level auditory facilitatory effect on visual processing.

In this experiment, we found that human voice can be integrated with invisible human face, and that is not due to low-level feature integration. Compared to the finding that a pure tone can enhance the processing of arbitrary visual stimulus under the CFS paradigm (Y.-H. Yang & Yeh, 2014), the current result shows unconscious multisensory integration at a semantic level, that is, face-voice pairing. Based on this observation, we adopted different facial expressions and paired them with congruent and incongruent affective voices to investigate unconscious multisensory binding with different emotions in the next experiment.

Experiment 2

The purpose of this experiment was to test whether the unconscious binding of face-voice can be specific to different emotions. Facial expressions (happy and fearful) and affective voices (laughing and screaming) were orthogonally paired to manipulate the emotional congruency between face and voice. Moreover, to better illustrate the unconscious multisensory binding under CFS, not only the b-CFS time but also eye-movements were recorded during CFS.

A leading study has adopted oculomotor response to measure unconscious visual processing under CFS (Rothkirch, Stein, Sekutowicz, & Sterzer, 2012). In their study, a Gabor patch was presented to one of six locations in a visual search task, and they found that participants had their eyes fixated longer at the location where the Gabor appeared (i.e., longer dwell time) than its mirror location. This result suggests that measurement of eye-movements is reliable in reflecting how the participants processed invisible information under interocular suppression. Here we incorporated two indices, saccade latency and dwell time, to provide further evidence of unconscious multisensory integration under CFS. Saccade latency refers to the time delay between the target onset and the saccadic eye-movement to the target location. Saccades are eye-movements that change from the current fixation to certain salient information, thus saccade latency could be used as an index for measuring how invisible facial expression attracts visual attention to its corresponding location under CFS if invisible stimuli indeed attract attention as shown by a previous study (Jiang, Costello, Fang, Huang, & He, 2006). Furthermore, we also included dwell time as a measure, which is calculated by the mean duration of gaze staying on the location of the face stimulus. If the congruency between facial expression and affective voice can be integrated unconsciously, measurements of dwell time and b-CFS time should reveal a similar pattern.

In addition, by pairing the same facial expression with different emotional prosodies, we can also examine whether it was facial expression (and not low-level features) that was processed during CFS. As the assumed "congruent" or "incongruent" face-voice pairs should be based on emotional contents (i.e., happy and angry) from both modalities, we predicted that invisible facial expressions would be detected faster with congruent emotional prosodies than incongruent ones only if the emotional meanings were extracted from facial expressions. Along the same line, we also predicted that low-level features of face (e.g., contrast, orientation, or spatial frequency) or neutral face that do not contain emotional information would not lead to the congruency effect.

Method

Participants. Another group of 20 participants with the same characteristics as described in Experiment 1 was recruited in this experiment.

Stimuli and apparatus. The experiment was controlled by Matlab (the Math Works, Natick, USA) with Psychophysics Toolbox 3.0 (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007). Participants viewed visual stimuli that were presented on a 19-inch i-TECH IF900 CRT monitor (1024 × 768 resolution at 120 Hz refresh rate) through active shutter glasses (NVIDIA 3D vision 2) for dichoptic viewing. The masks and face stimuli switched with the refresh rate of the monitor and shutter glasses, thus each eye viewed either the masks or the face stimuli. To record eye-movements, an Eyelink 2000 (SR Research Ltd., Canada) was used at 1000 Hz sampling rate.

To increase scanning area for eye-movements, the two square contours (23° × 23°) and the face stimuli (5.52° × 5.52°) were enlarged, and the face was equally presented at one of six possible positions (Rothkirch et al., 2012), which included one of three vertical areas (6.44° above, below, or on vertical center) at left or right visual field (4.6° from horizontal center). The other properties of face and masks were the same as those in Experiment 1. The face stimuli (6 posers, 3 females) were also generated by FaceGen (Singular Inversions Inc.) software, and each poser presented happy and fearful expressions. The voice stimuli (6 posers, 3 females) were adopted from the Montreal Affective Voices (MVA, Belin, Fillion-Bilodeau, & Gosselin, 2008), and each poser included laughing and screaming voices.

Design and procedure. We manipulated facial expression (happy and fearful) and congruency of face-voice pairing (congruent: happy-laughing/fearful-screaming pairs; incongruent: happy-screaming/fearful-laughing pairs) in a 2 x 2 within-subjects factorial design. The facial expressions and affective voices were orthogonally manipulated and the combinations were pseudo randomized within a run. There were 192 trials in total, divided into 4 runs. At the beginning of each run, a calibration of the eye tracker was conducted. Participants pressed “0” on the keyboard to start each trial, and were required to press “a” or “d” as soon as possible if they detected any part of a face that was presented either at the left or right visual field. After that, they were asked to judge the valence of affective voice that was either laughing or screaming by pressing the keys “5” or “9”, respectively.

Results

In this experiment, we not only measured RTs but also recorded eye-movements during the suppression state. There were two indices for the eye-movements data. First, we measured the latency from face onset time to the first saccade on the face location, which may reflect how efficiently a face attracted eye-movements to a specific location. Second, we measured the dwell time, the mean duration of fixation on the face location, which could be an index of how long that participants processed the congruency between face and voice. The accuracies of location detection (mean accuracy = 98.85%, SD = .78%) and affective voice judgment (mean accuracy = 96.53%, SD = 4.97%) were high.

RTs data. The correct RTs were submitted to a 2 (facial expression: happy and fearful) x 2 (congruency: congruent and incongruent) two-way repeated-measures ANOVA. The main effect of valence was significant [$F(1,19) = 6.803$, $MSE = .001$, $p = .017$, $\eta_{p}2 = .264$]; happy face ($M = 1.660$, $SD = .234$) was detected faster than fearful face ($M = 1.679$, $SD = .229$). The main effect of congruency was also significant [$F(1,19)$

= 6.027, *MSE* = .003, *p* = .024, $\eta_p 2$ = .241]; the congruent condition (*M* = 1.654, *SD* = .235) was detected faster than the incongruent condition (*M* = 1.685, *SD* = .227). There was no interaction between facial expression and congruency [*F*(1,19) = 1.985, *MSE* = .001, *p* = .175, $\eta_p 2$ = .095] (see Figure 2.3).

Figure 2.3. RTs data in Experiment 2. Happy face was overall detected faster than fearful face, and the congruent condition was detected faster than the incongruent condition.

Saccade latency. The latency of the first saccade on face location revealed significant main effect of facial expression [*F*(1,19) = 5.016, *MSE* = .004, *p* = .037, $\eta_p 2$ = .209]: happy face (*M* = .831, *SD* = .355) led to shorter saccade latency than fearful face (*M* = .862, *SD* = .353). There was nether main effect of congruency [*F*(1,19) = .655, *MSE* = .006, *p* = .428, $\eta_p 2$ = .033] nor interaction [*F*(1,19) = 2.345, *MSE* = .003, *p* = .142; $\eta_p 2$ = .110].

Dwell time. The dwell time of fixation on face location revealed significant main effect of congruency [*F*(1,19) = 4.745, *MSE* = .002, *p* = .042, $\eta_p 2$ = .200]; the congruent condition (*M* = 373, *SD* = .224) led to shorter dwell time than the incongruent condition (*M* = 373, *SD* = .224). There was neither main effect of valence [*F*(1,19) = .387, *MSE*

= .001, p = .541, η_p2 = .020] nor interaction [F(1,19) = 1.274, MSE = .001, p = .273, η_p2 = .063] (see Figure 2.4).

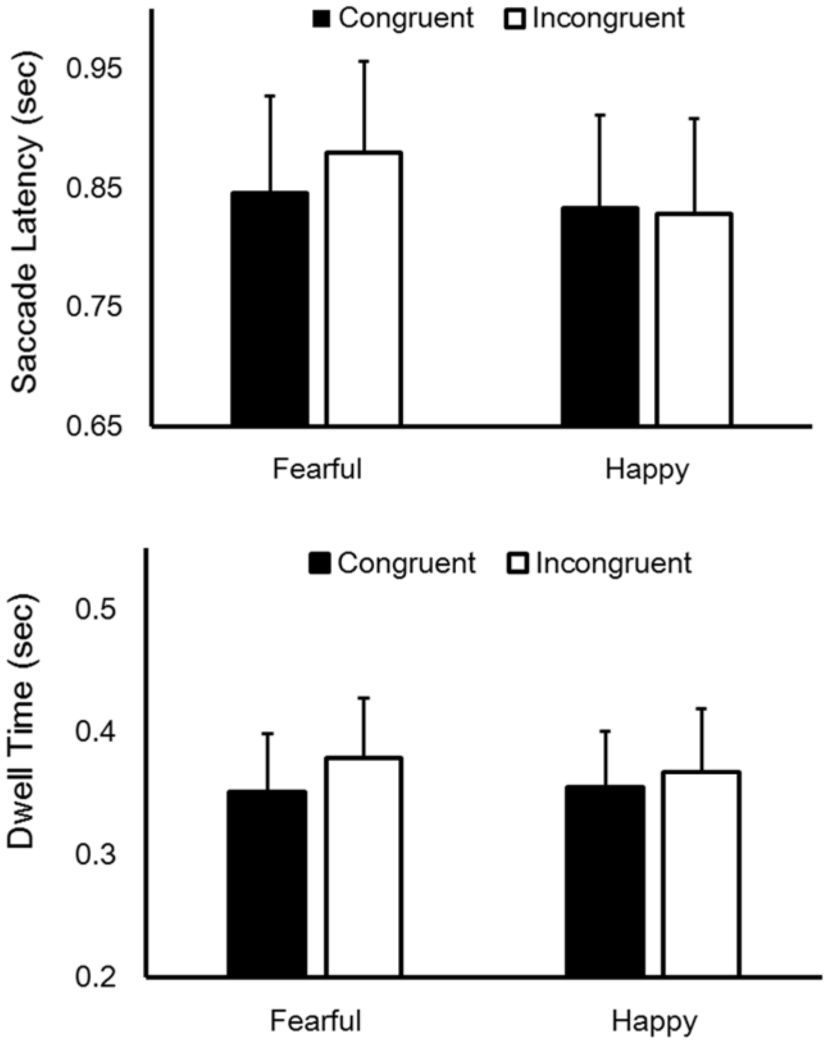

Figure 2.4. Eye-movement data in Experiment 2. Happy face revealed shorter saccade latency than fearful face (top panel), and the congruent condition revealed shorter dwell time than the incongruent condition (bottom panel).

Discussion

By measuring b-CFS time and eye-movements, this experiment examined whether face-voice binding can be specific to emotion under CFS. For the RTs data, we found that happy face was detected more quickly than fearful face, and congruent face-voice pairing was detected more quickly than incongruent pairing. In addition, the indices of eye-movements also corresponded to different effects in the RTs data. We found that first saccade latency on face location was correlated with the processing of facial expression. The shorter saccade latency on happy face than fearful face may reflect the efficiency of happy face to attract eye-movements during suppression, and consequently facilitated the happy face to release from suppression more quickly than fearful face. Moreover, we found that dwell time of fixation of face location was shorter in the congruent condition than the incongruent condition, which may also reflect multisensory enhancement such that a specific facial expression was processed faster with an emotionally congruent voice than with an incongruent one. The convergent evidences in RTs and eye-movements imply that emotional information laden by face and voice can be integrated unconsciously. Furthermore, since the detection of specific facial expression was facilitated by its corresponding affective voice, this also suggests that emotional information of facial expression can be extracted under the interocular suppression state.

Note that the saccade latency observed here was longer (around 850 ms) than traditional findings (around 200 ms, e.g., Q. Yang, Bucci, & Kapoula, 2002). This inconsistent observation might be due to the contrast manipulation of face and/or unconscious processing. To successfully induce interocular suppression at the beginning of each trial, a typical technique was to ramp up the contrast of the critical stimulus for preventing abrupt onset (Jiang et al., 2007; E. Yang et al., 2007). However, this gradual increase of face contrast may induce longer processing time for moving eyes to a given location than if a sharp contrast was implemented. More specifically, this saccade latency was measured under interocular suppression, which reflected unconscious detection of face. Compared to the visible target, participants may not be necessary to move their gaze to an invisible face. Although the saccade latency is a useful index to predict how efficiently an invisible target is detected, future works are needed to examine whether this relatively long latency was due to a specific processing under interocular suppression.

General Discussion

This study adopted the CFS paradigm to examine whether the binding of face and voice is processed unconsciously. Results from Experiments 1 and 2 showed that multisensory integration of emotional information could be processed under CFS. In Experiment 1, only voice (compared to beep or no-sound) facilitated the detection of an upright face. This result rules out the possibility that the observation is due to the low-level facilitatory effect of sound because the beep-sound did not facilitate the detection of upright-faces. This result was also not due to the inefficiency of beep-sound either because the beep-sound still facilitated the detection of inverted-faces, which is consistent with our previous finding that beep-sound can speed up arbitrary visual information under CFS (Y.-H. Yang & Yeh, 2014). Together with the face inversion effect (i.e., upright faces were detected faster than inverted ones) our findings support the idea that human voice has a specific role to integrate with human face when the orientation is familiar (i.e., upright). In Experiment 2, happy face yielded shorter RTs than fearful face, and consistent result was found in saccade latency on face location, which implies that unconscious processing of facial expression can attract one's fixation to a specific location even before the detection of face under interocular suppression. In addition, affective voices facilitated the detection of congruent emotional faces (i.e., happy-happy or angry-angry face-voice parings); this result was also supported by the dwell time of fixation on the face location (i.e., shorter dwell time in the congruent than the incongruent condition). Taken together, these results suggest that the binding of face-voice under CFS not only extracts facial information but also specific emotions.

One may argue that the sound-facilitation effect on face we have found here is a kind of priming effect in which sound primes the face if the sound and face were perceived at different times. We cannot directly exclude this possibility but we believe that our manipulations are different from other conventional priming experiments: The priming effect usually is based on relevant stimuli being presented sequentially, in which the prime more fluently paves the way for the processing of the target (Collins & Loftus, 1975; McNamara, 1992). However, we presented the sound at the same time as the onset of visual stimuli. In other words, the participants "received" the visual and auditory information simultaneously. In our case, although the visual information is not consciously "perceived", it is still being continuously "received" and processed in the brain, which integrated with the auditory stimulus at the same time. Presenting multisensory stimuli simultaneously is different from those studies that use sequential presentation of stimuli to probe the priming effect under CFS (e.g., Bahrami et al., 2010; Costello et al., 2009). As a stimulus being processed in the neural pathway and being represented for conscious perception might differ fundamentally, it is important to discriminate the time course of multisensory integration of visual and auditory emotional information in sensation and (awareness) perception (cf. Pourtois, de Gelder, Vroomen, Rossion, & Crommelinck, 2000).

Our results support previous lesion studies that involved participants with loss of visual awareness and probed their unconscious multisensory integration of emotional information. Compared to the debates that lesion studies may involve functional recovery (Mogensen, 2011) and/or residual visual experiences (See Overgaard, 2011 for discussion), the current study provide new evidence for unconscious binding of multisensory emotional information in neurological intact participants. Some neuroimaging studies seem to hint at possible neural loci for this unconscious multisensory integration of emotional information. For example, certain neural correlate regions such as the middle temporal gyrus (MTG), superior temporal sulcus (STS) and amygdala are responsible for conscious face-voice integration (Kreifelts, Ethofer, Grodd, Erb, & Wildgruber, 2007; Kreifelts, Ethofer, Shiozawa, Grodd, & Wildgruber, 2009; Pourtois, de Gelder, Bol, & Crommelinck, 2005). On the other hand, unconscious uni-sensory processing of facial expression activates the STS, amygdala, pulvinar, and insula, etc. (Jiang & He, 2006; Tamietto & de Gelder, 2010). Based on the coincident neural mechanisms such as STS and the amygdala, between consciously multisensory integration and unconsciously unisensory facial processing, we propose that these areas might also be the neural loci for unconscious binding of multisensory emotional information. Further studies using similar settings as the current study could examine this possibility directly.

Compared to earlier studies, the current study provides new evidence that facial expression can be processed under CFS. Had only facial expression been probed under the CFS alone, it would be inconclusive as to whether it is the facial expression or some low-level features that had contributed to the results. Because there are always some low-level features embedded in facial expressions and covary with them (Gray et al., 2013; T. Stein & Sterzer, 2012), the twofold explanations cannot discriminate as to which level of processing the probed stimuli has gone through under the CFS. In this regard, the current study provides new evidence and methodological rationale to resolve this issue. Instead of trying to keep the entire low-level feature consistent (if this is possible), we manipulated different kinds of sound paired with a face, that is, face-voice pairing in Experiment 1 and facial expression- affective voice pairing in Experiment 2, to investigate the unconscious processing of facial information. These results are based on the same visual features but accompanied with different sounds, thus different release times from suppression revealed that they are not caused by the different low-level features embedded in the face but rather due to extracting the high-level information. From the evidence of multisensory integration of face-voice binding at different levels (i.e., facial and emotional information here), we conclude that a face can be emotionally processed at the level we manipulated.

The coherent result of eye-movement measurement with the b-CFS time also provides two implications on methodological issue of CFS in the study of unconscious processing (see E. Yang, Brascamp, Kang, & Blake, 2014 for a review). First, recording eye-movements per se is a suitable measurement to reveal different processing under CFS. As mentioned previously, previous study showed that an invisible Gabor can trigger longer dwell time for the Gabor location than its counterpart (Rothkirch et al., 2012). Here we further took steps to show that dwell time can be an index to directly differentiate the congruency of multisensory binding (i.e., congruent pair had shorter dwell time than incongruent pair). Moreover, we adopted saccade latency as another index, which also revealed different times on the processing of facial expression (i.e., happy face had shorter latency than fearful face). These results suggest that recording eye-movements is an advanced and feasible measurement under CFS for future studies.

Second, our eye-movement results also support that the b-CFS time is a reliable measurement. Although the b-CFS time was a well adopted measurement to probe how efficiently some information being processed under CFS (see Gayet et al., 2014 for a review), recent studies argued against the b-CFS time: it may reflect not only CFS-specific processing but also involve some conscious processing *after* CFS (T. Stein, Hebart, & Sterzer, 2011; T. Stein & Sterzer, 2014). Here our eye-movement result was recorded during CFS, and both saccade latency and dwell time occurred *before* the b-CFS time. Nevertheless, the pattern of b-CFS time in different facial expression and face-voice congruency corresponded with the eye-movement data. These results imply that the b-CFS time can be entirely explained by the unconscious processing (i.e., eye-movement) preceding the detection of the face stimulus, thus we conclude that b-CFS time is an index more reflecting of CFS-specific processing rather than the processing after CFS.

In conclusion, by adopting the CFS paradigm with suppression time measurement and eye-movement recording, this study demonstrates unconscious binding of facial expression and affective voice in neurologically intact participants. This finding suggests that although understanding other's emotion via integration of distinct modalities is critical in social interaction, consciousness is not critical to form this unified percept.

Study 3

Unconscious Processing of Emotional Information as Revealed by Priming under Interocular Suppression

(Submitted)

Abstract

Whether emotional information can be processed unconsciously is still controversial; the debate is partially due to an unclear boundary between consciousness suppression and non-suppression state, and ambiguous low-level property contributions. We tackled this issue by adopting an affective priming paradigm with the continuous flash suppression method to render an emotional prime invisible. Following the invisible prime, a visible target with emotionally congruent or incongruent valence was presented, and emotional judgment of the target was conducted. The priming effect obtained varied with the type of prime (face or word). Both negative and positive faces revealed a positive priming effect. However, only positive word showed this pattern while negative word revealed a negative priming effect, which could be due to rich connotations embedded in negative words relative to positive words. Since suppressed prime can induce affective priming effect on subsequent target, these results suggest that emotional information can be processed unconsciously.

Introduction

When different visual stimuli are presented to each of the two eyes simultaneously, the incompatible information in the two eyes would rival for gaining access to visual awareness (Alais & Blake, 2005). As a consequence, an observer tends to perceive only one of the two stimuli, with the other stimulus being interocularly suppressed from awareness. It has been proposed that this reciprocal competition occurs at the early stages of visual processing (Blake & Logothetis, 2002; Polonsky, Blake, Braun, & Heeger, 2000; Sengpiel, Blakemore, & Harrad, 1995), and whether the suppressed information can still be processed along the information processing pathway depends on the level of processing the stimulus requires.

Most findings suggest that low-level but not high-level information can be processed under interocular suppression. For example, low-level properties of suppressed image such as spatial frequency (Blake & Fox, 1974), orientation (Wade & Wenderoth, 1978) and motion direction (Lehmkuhle & Fox, 1975) can induce aftereffect on subsequent target. On the other hand, high-level information such as object representation (Tong et al., 1998) and semantic information (Blake, 1988; Zimba & Blake, 1983) does not seem to survive interocular suppression. Zimba and Blake (1983) adopted the semantic priming paradigm with binocular rivalry and found that priming effect occurred only when the prime was visible (i.e., in the dominant phase) and not when it was invisible (i.e., in the suppressed phase). Consistent with this result, both monkey electrophysiology (Sheinberg & Logothetis, 1997) and human *f*MRI (Tong et al., 1998) showed that activities of neural system are correlated with conscious perception of the image; but for image that is not consciously perceived during binocular rivalry, there seems to be no association with neural activity. These findings contributed to the proposition that high-level information under suppression is "out of sight, out of mind" (Zimba & Blake, 1983).

In the last decade, Tsuchiya and Koch (2005) developed another type of interocular suppression technique called the *continuous flash suppression* (CFS), which provides deeper suppression and yet longer processing time of the stimulus (if it gets processed) than conventional binocular rivalry (Tsuchiya et al., 2006). In this paradigm, the to-be-tested stimulus is presented to one eye, and is strongly suppressed by a train of dynamically high-contrast masks that are presented to the other eye. By manipulating the contrast between stimuli and masks, the suppression time could be well controlled for sustaining from seconds to minutes (Tsuchiya & Koch, 2005). Since invisible information requires duration long enough to process (Lo & Yeh, 2008), this method provides an opportunity for information to be fully processed under interocular suppression. Indeed, Jiang et al. (2007) manipulated the familiarity of face orientation and the recognizability of word, and found that upright face and recognizable word have faster time of breaking CFS (b-CFS) than their unfamiliar counterparts.

Several studies used the b-CFS paradigm to test whether emotional information of facial expression can be processed under CFS (T. Stein & Sterzer, 2012; Sterzer et al., 2011; E. Yang et al., 2007). For example, E. Yang et al. (2007) showed that fearful face gained *faster* time of b-CFS compared to neutral or happy face; and a similar pattern was found when the orientation of the face was inverted (Gray et al., 2013; E. Yang et al., 2007), as well as for an observer with lesioned amygdala (Tsuchiya et al., 2009). However, regarding the emotional processing under CFS, the emotion information laden by word revealed a very different pattern to that by facial expression. For example, Y. H. Yang and Yeh (2011) adopted negative emotional Chinese words that are either emotion descriptive (e.g., fearful) or emotion inducing (e.g., abuse), to compare with neutral words in the upright and inverted orientation. They not only found that negative words consistently had *slower* time of b-CFS than neutral ones (see also Prioli & Kahan, 2015 for English word), but this pattern also vanished with word inversion. Due to the contradictory results between facial expression (i.e., negative priorities) and emotional word (i.e., negative impairment) from previous studies, whether prior b-CFS studies probed similar high-level processing of different types of stimuli with emotional content should be the next question (see Gayet et al., 2014 for a review).

Explanation of findings indexed by the measure of b-CFS time is still an issue for the following reasons. First, since b-CFS time actually measures how long it takes for a participant to respond to stimulus detection on top of the duration under CFS, the assumption that b-CFS time can be fully attributed to the unconscious processing *during* CFS is debatable; that is, the time interval between stimulus detection and response is also included in b-CFS time, and thus differences in that interval (which is not unconscious processing) may have contributed to different b-CFS times (Jiang et al., 2007). To address this issue, Jiang et al. (2007) adopted a binocular non-suppression condition, which superimposed the to-be-tested stimuli on top of the dynamic masks to mimic the percept during the interval between stimulus detection and response. If no difference was found in this non-suppression condition, then the significant results from b-CFS time can be attributed to the CFS-specific unconscious processing during suppression (e.g., Gayet, Paffen, & Van der Stigchel, 2013; Jiang et al., 2007; Mudrik, Breska, Lamy, & Deouell, 2011; Y. H. Yang & Yeh, 2011). However, having the non-suppression control as a comparable condition was argued against since the distribution of reaction times (RTs) in this condition differed too much from that in the CFS condition (T. Stein, Hebart, et al., 2011). In summary, these observations raised an issue that b-CFS time may not purely reflect CFS-specific unconscious processing (T. Stein & Sterzer, 2014).

The second issue was that since b-CFS time is a direct measure which tests how efficient a stimulus can be processed under CFS, if different classes of property are contained within a stimulus, it is hard to distinguish which class contributed to the result. For instance, in earlier studies that found preferential access of fearful face into awareness, another critical factor was the salient eye-white region embedded in fearful face (e.g., E. Yang et al., 2007). While one could explain that eye-white region is representing the *emotional valence* of fearful face (Whalen et al., 2004; E. Yang et al., 2007), others could argue that eye-white region only reflects *low-level properties* such as strong contrast in subareas (Gray et al., 2013). Based on these considerations of b-CFS, it is important to use other methods to clarify which level of unconscious processing is probed under CFS (see E. Yang et al., 2014 for a review).

The purpose of the current study is twofold. First, we will examine whether emotional content can be extracted under CFS by adopting the priming paradigm: an emotional prime is rendered invisible under CFS, and the RTs of emotional judgment of subsequent emotional target is measured. If the *emotional content* of the invisible prime can be extracted, then the congruency of emotional valence between the prime and the target would affect the emotional judgment of the target. By adopting this indirect measure, the results can provide evidence for high-level emotional specific extraction under CFS without including low-level information such as contrast (Faivre, Berthet, & Kouider, 2014). Second, we aimed to compare the priming effect between word and face under CFS. As mentioned previously, whereas fearful face had faster b-CFS time than neutral faces (Tsuchiya et al., 2009; E. Yang et al., 2007), negative words had slower b-CFS time than neutral words (Prioli & Kahan, 2015; Y. H. Yang & Yeh, 2011), Since these studies all claimed that emotional content can be processed under CFS, it is important to learn whether word and face induce similar priming effect under CFS in the same experimental setting.

Experiment 1

The purpose of the first experiment was to test whether the emotional meanings of face and word can be extracted under CFS. More specifically, by adopting the priming paradigm, we tested whether a face prime (Experiment 1a) or a word prime (Experiment 1b) with emotional content, which was suppressed by a series of high contrast masks for 1000 ms, can affect the efficiency of emotional judgment on the subsequently presented target. Importantly, to prevent the possibility that any priming effect was due to similar perceptual properties between the prime and the target (Faivre et al., 2014), a face and a word were mutually manipulated as prime and target respectively in Experiment 1a and vice versa in 1b. Since faces and words have minimal overlapping features, possibilities of perceptual facilitation can be ruled out if priming effect was found. In addition, to better ensure that the participants were fully unaware of the emotional prime under CFS, we excluded trials with primes that were seen by asking each participant to report subjective percept, and to perform an objective visibility check conducted after the main priming task.

Method

Participants. Twenty participants took part in this experiment. They were naïve about the purpose of this experiment and had normal or corrected-to-normal vision. All participants gave informed consent according to the rules of the Ethics Committee of the Department of Psychology, National Taiwan University.

Stimuli materials and apparatus. Stimuli were controlled by Psychophysics Toolbox 3.0 (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007) on Matlab 2012a (the Math

Works, Natick, USA) and were presented on a 19-inch i-TECH IF900 CRT monitor (1024 × 768 resolution at 120Hz refresh rate). To present stimuli dichoptically, observes viewed the stimuli through a pair of active shutter glasses (NVIDIA 3D vision 2) which were synchronized with the refresh rate of the monitor via infrared emitter (NVIDIA 3D vision 2), thus each eye viewed either the Mondrian-masks or the prime stimuli. The masks and prime stimuli also refreshed with the same rate as the monitor.

Two square contours (13.8° × 13.8°, 0.18° thickness with white noise pattern) were presented on a gray background (RGB values of [127, 127, 127]) for stable dichoptic fusion. A Mondrian-like masks with 3000 patches (random size: 0.02° to 1.09°; random gray scale: 0~255) was presented in one fusion contour and a face prime (7.36° × 7.36°; Experiment 1a) or a word prime (5.52° × 2.76°; Experiment 1b) was presented in another fusion contour. To induce strong interocular suppression, the masks were always presented to the participant's dominant eye, and the prime stimuli were always presented to the non-dominant eye. The word target (5.52° × 2.76°; Experiment 1a) or face target (7.36° × 7.36°; Experiment 1b) was presented binocularly in the center of both fusion contours without any masks.

The face stimuli (12 posers; 6 females) contained happy and fearful facial expressions that were adopted from Ekman and Friesen (1976), and the word stimuli contained positive (12 words) and negative (12 words) meanings that were matched in usage frequency [mean frequency per one million words: positive, 28.75; negative, 29.69, $paired\text{-}t(11) = -.283$, $p = .783$] and stroke numbers [mean stroke counts: positive, 24.92; negative, 26.00, $paired\text{-}t(11) = -.816$, $p = .432$]. During the priming-CFS task, to prevent the abrupt stimulus onset from triggering prime perception, the contrasts of face prime and word prime were ramped up gradually from 0% to 25% within 1000 ms. However, if the participants reported that they saw the prime during CFS after a trial, the contrast would reduce 1.5% to make the prime less visible in later trials. The contrast of masks was kept at 100% throughout the experiment.

Design. A within-subjects experiment with factors emotional valence of prime (positive and negative) and prime-target congruency (congruent and incongruent) was conducted. The emotional valence, congruency, and prime identity were orthogonally manipulated and repeated four times with pseudo randomization. There were 192 trials in Experiment 1a (Face-prime with word-target) and Experiment 1b (word-prime with face-target), respectively, and the order of Experiment 1a and 1b was counterbalanced between participants.

Procedure. In the main priming task, the participants pressed the ‘z’ key to start every trial, and a 300 ms blank was presented for preparation. Following the blank was the CFS stage for 1000 ms, and a prime with either positive or negative meaning was suppressed by the dynamic masks that flashed at 10 Hz. After CFS, the target with either positive or negative meaning was presented for 100 ms. The participants were asked to judge the emotional valence of the target by pressing ‘5’ for positive and ‘6’ for negative meaning. RTs of the emotional judgment to the target served as an index for the priming effect. In addition, if the participants saw any parts of prime during CFS, they were instructed to press ‘7’ to abandon the trial, which served as a subjective criterion of invisible prime (see Figure 3.1).

After the primed emotional judgment task, participants performed an objective visibility check task with 48 trials, with equal probability for positive and negative valences. The procedure was exactly the same as the primed task except without target stimulus. The contrast of stimuli always ramped up gradually from 0% to 25% within 1000 ms. The participants were asked to judge the emotion valence (positive or negative), in a two-alternative forced choice procedure, of the invisible prime during CFS by pressing ‘5’ for positive and ‘6’ for negative valence.

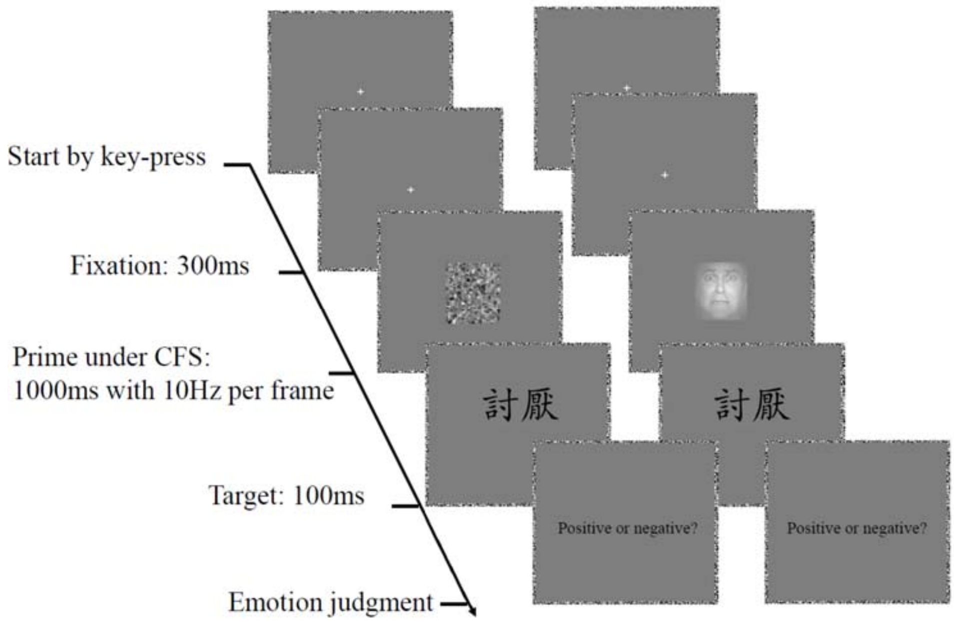

Figure 3.1. Stimuli and procedure. Trial started by key pressing, followed by a fixation cross presented for 300 ms. After this, a prime stimulus (shown here as a fearful face on the right side) was suppressed by dynamic masks that refreshed every 100 ms. Then, a target (Chinese two character words, “hate” in this example) was presented for 100 ms, and the participants were asked to judge the emotional valence of the target. They were instructed to press another key to abandon this trial if they saw any parts of the prime during

Results

Experiment 1a (face prime with word target). One participant was excluded because of high accuracy in objective visibility check (78.50%). Besides this, the mean accuracy of objective visibility check was at chance level [M = 50.26%, SD = 4.06%, $t(18)$ = .27, p = .79]. In addition, the rate of seeing the face prime during CFS was low (M = .68

%, *SD* = 1.54 %) and the accuracy of emotional judgment of word target was high (*M* = 96.41%, *SD* = 2.64%).

RTs of correct emotional judgment were submitted to a two-way repeated measure analysis of variance (ANOVA) with the factors valence (positive and negative) and congruency (congruent and incongruent). The main effect of congruency was significant [$F(1,18) = 7.10$, $MSE = 339.25$, $p = .016$, $\eta_p^2 = .28$], reflecting shorter RTs in the congruent condition (*M* = 592.26 ms, *SD* = 106.68 ms) than the incongruent condition (*M* = 603.52 ms, *SD* = 109.82 ms). There was neither significant main effect of valence [$F(1,18) = .02$, $MSE = 706.92$, $p = .90$, $\eta_p^2 < .01$] nor significant interaction [$F(1,18) = .01$, $MSE = 1729.56$, $p = .94$, $\eta_p^2 < .01$] (see Figure 3.2a).

Experiment 1b (word prime with face target). Two participants were excluded for further analysis; one was due to high probability of seeing the prime during CFS (66.67 %) and the other was due to high accuracy in objective visibility check (77.15%). The rest of the data showed that the rate of seeing the prime during CFS was low (*M* = 1.23%, *SD* = 3.04 %), the accuracy of visibility unawareness check was at chance level [*M* = 51.98 %, *SD* = 9.07%, $t(17) = .93$, $p = .37$], and the accuracy of emotional judgment of target was high (*M* = 93.20 %, *SD* = 6.00 %).

For ANOVA with RTs from correct trials, there was a significant interaction between valence and congruency [$F(1,17) = 5.66$, $MSE = 3237.04$, $p = .03$, $\eta_p^2 = .25$]. In the positive prime condition, the simple main effect of congruency [$F(1,34) = 3.38$, $MSE = 1994.86$, $p = .07$, $\eta_p^2 = .09$] showed marginal significance that the congruent condition (*M* = 658.00 ms, *SD* = 165.34 ms) was judged faster than the incongruent condition (M = 685.38 ms, *SD* = 170.78 ms). On the other hand, the simple main effect of congruency in the negative prime condition [$F(1,34) = 5.92$, $MSE = 1994.86$, $p = .02$, $\eta_p^2 = .15$] revealed that the incongruent condition (M = 653.71 ms., *SD* = 175.51 ms) was judged faster than the congruent condition (M = 689.93 ms, SD = 170.11 ms). There was neither significant main effect of valence [$F(1,17) < .01$, $MSE = 630.06$, $p = .98$, $\eta_p^2 < .01$] nor significant main effect of congruency [$F(1,17) = .47$, $MSE = 725.67$, $p = .50$, $\eta_p^2 = .03$] (see Figure 3.2b).

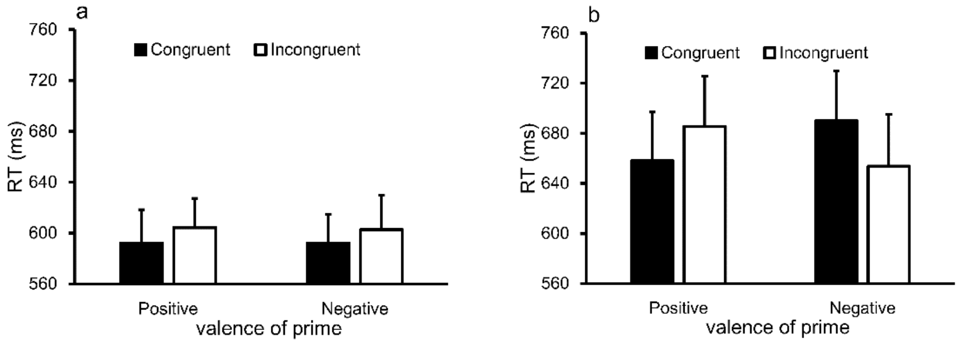

Figure 3.2. Mean RTs in Experiment 1a (Figure 3.2a) and Experiment 1b (Figure 3.2b).

Discussion

We found the unconscious semantic priming effect in this experiment; that is, the affective meaning of invisible prime under CFS can affect the emotional judgment speed of subsequent target. In experiment 1a, the emotional face prime induced positive priming effect, which revealed that congruent meaning between prime and target can facilitate emotional judgment, and this result was unchanged with different emotional valences of prime. On the other hand, in experiment 1b, we also found that emotional word prime could affect emotional judgment speed of subsequent target, but it depended on the emotional valence of prime. That is, the positive word prime and negative word prime induced positive priming and negative priming, respectively. The results from experiment 1a and 1b cannot be simply explained by perceptual priming since the prime and target used here did not share similar perceptual properties. In addition, the prime during CFS was fully suppressed as supported by subjective percept and objective visibility check. Based on the finding that congruency between invisible prime and target can modulate emotional judgment efficiency, these evidences imply that emotional meaning of face and word can be extracted under CFS.

Given that negative priming was found in negative word prime condition, it is important to know the mechanism for this reversed pattern. One possible explanation is from the neural habituation priming model (Huber, 2008; Huber & O'Reilly, 2003), which suggests that the neural system would become habituated to old information so as to increase the sensitivity to deal with new stimulus. The consequence of habituation would transfer the priming effect from facilitation to impairment if the increase in prime duration induced 'cognitive aftereffect'. This idea has been used to explain why words with negative valence were detected more slowly than neutral ones under CFS: negative words triggered stronger neural response and consequently greater habituation compared to the neural word (Huber, 2015; Prioli & Kahan, 2015). Regarding the simulation result of the negative word in CFS by the neural habituation model (Huber, 2015), the 1000 ms prime duration used in this experiment was likely the cause of the negative priming effect. We therefore used shorter prime duration (200 ms) in the next experiment to test this possibility.

Experiment 2

To test whether the negative priming effect found in the negative word prime condition was due to long prime exposure that induced cognitive aftereffect, we shortened the prime duration to 200 ms in this experiment. The same design and procedure as in Experiment 1 was used for comparison between the two experiments.

Method

Participants. Another group of 22 adults were recruited to participate with the same procedure as the previous experiment.

Stimulus materials, apparatus, design, procedure. All were kept the same as in Experiment 1, except that the prime duration was reduced to 200 ms. The contrast of the prime also ramped up from 0% to 25% within the 200 ms of the CFS stage.

Results

Experiment 2a (face prime with word target). During the CFS, all participants reported that they did not see the prime (M = 0, SD = 0), and the accuracy of objective visibility check was at chance level [*M* = 48.38 %, *SD* = 6.77%, *t*(21) = -1.10, *p* = .28]. The accuracy of emotional judgment of target was high (*M* = 95.43 %, *SD* = 4.11 %).

The RTs of correct emotional judgment of target showed a significant main effect of congruency [*F*(1,21) = 5.36, *MSE* = 939.99, *p* = .03, η_p^2 = .20], reflecting shorter RTs in the congruent condition (M = 491.68 ms, *SD* = 94.68 ms) than the incongruent condition (*M* = 506.81 ms, *SD* = 113.88 ms). There was neither significant main effect of valence [*F*(1,21) = .05, *MSE* = 935.77, *p* = .83, $\eta_p^2 < .01$] nor significant interaction [*F*(1,21) = .08, *MSE* = 2730.84, *p* = .40, η_p^2 = .03] (see Figure 3.3a).

Experiment 2b (word prime with face target). Five participants were excluded from further analysis because of high accuracy in objective visibility check (*M* = 79.83 %, *SD* = 8.35 %). The rest of the data showed low visibility of prime during CFS (*M* = .03 %, *SD* = .12%), and chance level for the accuracy of objective visibility check [*M* = 52.24 %, *SD* = 5.69%, *t*(16) = 1.58, *p* = .13]. The mean accuracy of emotional judgment of target was high (91.70%, *SD* = 5.83%).

Correct RTs showed that neither main effect of valence [*F*(1,16) = 1.10, *MSE* = 503.94, *p* = .31, η_p^2 = .06], nor main effect of congruency[*F*(1,16) = .03, *MSE* = 1136.67, *p* = .86, $\eta_p^2 < .01$], nor interaction [*F*(1,16) = .729, *MSE* = 14788.06, *p* = .41, η_p^2 = .04] was significant (see Figure 3.3b).

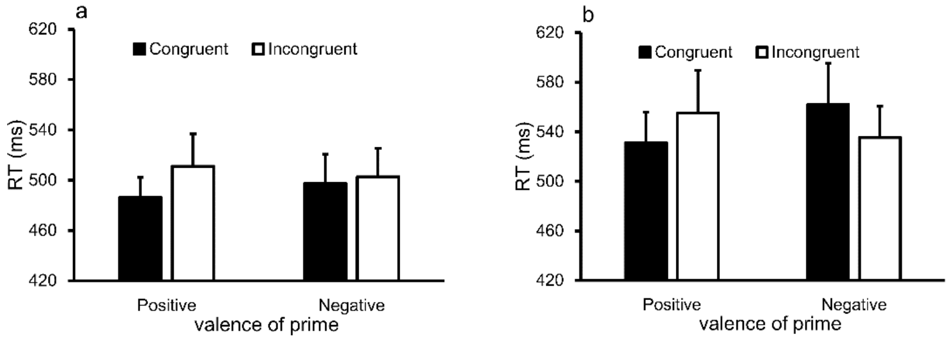

Figure 3.3. Mean RTs in Experiment 2a (Figure 3.3a) and Experiment 2b (Figure 3.3b).

Discussion

In this experiment, we used a markedly shorter prime duration (200 ms) relative to that in Experiment 1 (1000 ms). For the face prime, we replicated previous finding that the congruent context between invisible prime and target can facilitate the emotional judgment of target. On the other hand, the priming effect for the word prime reserved similar pattern but was not significant with shorter prime duration.

Although some studies showed that positive priming effect can become negative with lengthened prime duration under CFS (Barbot & Kouider, 2012; also see Faivre & Kouider, 2011a for similar conclusion in crowding), it was not found in the current study. In Barbot and Kouider's (2012) study, an invisible famous face under CFS served as prime, and the face identification of subsequent face target was facilitated by the same face at short duration (60 ms), but was impaired by the same face at long duration (1000 ms). However, note that the congruent condition of prime and target used in their study only differed in size and viewpoint but kept the face identity the same; their results therefore could be partially attributed by perceptual priming (Faivre et al., 2014). Since our experimental setting used very different types of prime and target, deeper processing up to the semantic level was needed. Therefore, instead of transferring from negative priming to positive priming with shorter duration of prime presentation, the priming effect was overall eliminated.

According to the neural habituation modeling of the negative word in CFS (Huber, 2015), there are three stages of semantic processing with the function of prime duration: no semantic activation at very short duration, positive priming at medium duration, and negative priming at long duration. While one may argue that 200 ms prime duration was not beyond the critical time point to reveal the positive priming effect in negative word, we excluded this possibility with the following reasons. First, if 200 ms was too short to induce a positive priming effect, a flat pattern or a weak positive priming pattern should be seen. However, we did not observe any tendency of this pattern translation here; instead, we still found a pattern of negative priming (though no significant; see Figure 3.3b). Second, if 200 ms was too long and was still at the negative priming stage, then an even shorter prime duration (<200 ms) should be used. However, based on the neural habituation model (Huber, 2015) and some empirical finding (Kang, Blake, & Woodman, 2011), the prime duration shorter than 200 ms would be too weak to generate any semantic activation. Therefore, the negative priming effect for negative words (Experiment 1b) could not be attributed to the 'cognitive aftereffect' account.

Another possibility that could explain the word priming effect in Experiment 1b could be that the emotional meaning of word prime could not be extracted in CFS, and the different patterns of priming effect between positive word prime and negative word prime were simply due to different responses to the face target. That is, in the positive word prime condition, the congruent (positive) face target was judged faster than incongruent (negative) face target; similarly, in the negative word prime condition, the incongruent (positive) face target was judged faster than congruent (negative) face target. In other words, the result from experiment 1b could be explained by the idea that the positive face targets were always judged faster than the negative face targets. This idea was supported by some studies that asserted emotional judgment of positive face was faster and more accurate than negative face (e.g., Calvo & Beltran, 2013; Tottenham et al., 2009). Since the results of experiment 1a and 2a showed that emotional judgment efficiency of word target significantly depended on the congruent prime rather than the response to the word target alone, we adopted only word targets to test whether the word priming effect can still be observed as in Experiment 1b.

Experiment 3

In this experiment, we took out the face targets and adopted only the word targets to test whether the positive and negative priming effect can still be observed from positive word prime and negative word prime, respectively. If the result of Experiment 1b was due to the influence of word prime, then similar priming pattern should be observed regardless which kind of targets were used. Alternatively, if different priming effect simply reflected the response bias to different face targets (i.e., positive face was judged faster than negative face), this pattern should be absent since face target was switched to word target here. Moreover, to better support previous results in experiment 1b and 2b, we also included long duration (1000 ms) and short duration (200 ms) as prime duration in experiment 3a and 3b, respectively.

Method

Participants. Another 46 participants were recruited in this experiment. Twenty-three participants contributed to Experiment 3a and 23 participants contributed to Experiment 3b.

Stimulus materials, apparatus, design, procedure. The method of Experiment 3a and 3b was similar to Experiment 1b and 2b in that both used emotional word as prime, and the prime duration was 1000 ms and 200 ms, respectively. However, the face target in previous experiments was changed to word target in this experiment. Another set of positive and negative words were selected as the word target, the usage frequency [mean frequency per one million words: positive, 58.60; negative, 61.27, *paired-t*(11) = -.096 ,*p* = .925] and stroke numbers [mean stroke counts: positive, 23.17; negative, 23.83, *paired-t*(11) = -.298 ,*p* = .772] were matched. To prevent perceptual overlapping between prime and target (Barbot & Kouider, 2012), the target word was larger (7.36° × 3.68°) than the prime word (5.52° × 2.76°).

Results

Experiment 3a (1000 ms of word prime with word target). Four participants were excluded from further analysis: two were due to high performance in objective visibility check (*M* = 76.47 %, *SD* = 16.05 %), one was due to high probability of seeing the prime during CFS (*M* = 12.50 %) and one quitted the experiment after feeling uncomfortable during CFS. The rest of the data revealed low visibility of prime during CFS (*M* = .11 %, *SD* = .36 %) and chance level performance in the objective visibility check [*M* = 51.37 %, *SD* = 4.50 %, *t*(18) = 1.29, *p* = .212]. The accuracy of emotional judgment of target was high (*M* = 96.41 %, *SD* = 2.40 %).

ANOVA on RTs with correct responses revealed significant interaction between emotional valence of prime and congruency [*F*(1,18) = 8.64, *MSE* = 2665.08, *p* = .01, η_p^2 = .32]. In the negative prime condition, simple main effect of congruency [*F*(1,36) = 8.23, *MSE* = 2371.58, *p* = .01, η_p^2 = .19] revealed that trials in the incongruent condition (M = 519.50 ms, SD = 130.90 ms) were judged faster than those in the congruent condition (M = 567.48 ms, SD = 197.93 ms). On the other hand, simple main effect of congruency in the positive prime condition only revealed a tendency [*F*(1,36) = 2.37, *MSE* = 2371.58, *p* = .14, η_p^2 = .06] that the congruent condition (*M* = 538.43 ms, *SD* = 162.29 ms) was judged faster than the incongruent condition (*M* = 563.41 ms, *SD* = 155.58 ms). There was neither main effect of valence [$F(1,18) = .78$, $MSE = 1105.04$, $p = .39$, $\eta_p^2 = .04$] nor main effect of congruency [$F(1,18) = 1.01$, $MSE = 2078.07$, $p = .33$, $\eta_p^2 = .05$] (see Figure 3.4a).

Experiment 3b (200 ms of word prime with word target). Four participants were excluded from further analysis; three were due to high accuracy in the objective visibility check (*M* = 60.60 %, *SD* = .87 %), and one was due to high probability of seeing the prime during CFS (*M* = 14.53 %). The rest of data showed that low visibility of prime during CFS (*M* = .85 %, *SD* = 1.77 %), and chance level performance in the objective visibility check [*M* = 51.67 %, *SD* = 3.78 %, $t(19) = 1.877$, $p = .08$]. The accuracy of emotional judgment of target was high (M = 94.13 %, SD = 3.05 %).

Correct RTs with ANOVA showed neither main effect of prime valence [$F(1,18) = .90$, $MSE = 699.41$, $p = .35$, $\eta_p^2 = .05$], nor main effect of congruency [$F(1,18) = 2.26$, $MSE = 1409.28$, $p = .15$, $\eta_p^2 = .11$]. The interaction between prime valence and congruency revealed a similar pattern as Experiment 3a, but did not reach significance [$F(1,18) = 2.95$, $MSE = 1985.67$, $p = .10$, $\eta_p^2 = .14$] (see Figure 3.4b).

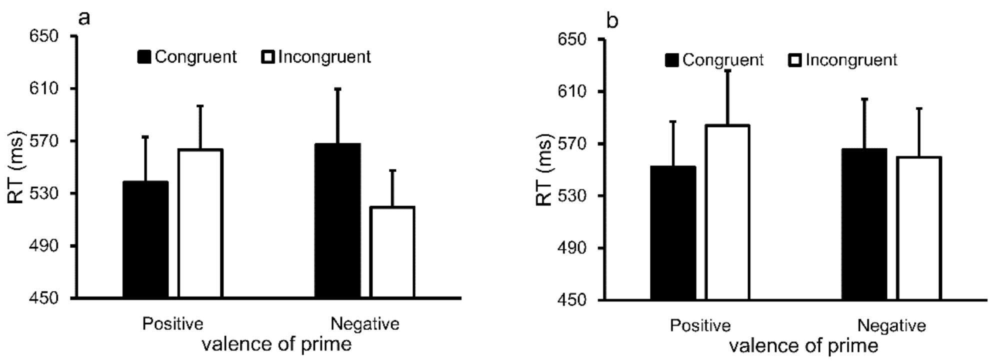

Figure 3.4. Mean RTs in Experiment 3a (Figure 3.4a) and Experiment 3b (Figure 3.4b).

Discussion

In this experiment, we adopted emotional words for both prime and target, and found that congruency between the prime and the target sill had an effect on the emotional judgment efficiency in long prime duration (Experiment 3a, 1000 ms), but this effect became weaker in short prime duration (Experiment 3b, 200 ms). Regarding prime valence, we found that congruent target was judged faster than incongruent target in the positive prime condition, and incongruent target was judged faster than congruent target in the negative prime condition. These findings were consistent with previous results with face target in Experiment 1b and 2b. Moreover, since the target was swapped from face to word in this experiment, the different priming pattern between positive prime and negative prime cannot be simply attributed to the different emotional judgment speed of face target.

General Discussion

In this study, we tested whether emotional content can be extracted under interocular suppression by adopting the affective priming paradigm with the CFS paradigm. The results revealed very different priming patterns between face prime and word prime. For the face prime, congruent meaning of prime facilitated emotional judgment of subsequent target. This effect was found in both positive face prime and negative face prime, across long priming duration (1000 ms, Experiment 1a) and short priming duration (200 ms, Experiment 2a). On the other hand, the word prime induced different priming effect depending on the emotional valence of prime. That is, the emotional judgment of target can be facilitated by the congruent meaning of positive word prime, but was impaired by the congruent meaning of negative word prime. This result was significant when the face target (Experiment 1b) or the word target (Experiment 3a) was used. However, this priming effect was dependent on prime duration, which was only observed in long duration (1000 ms, Experiment 1b and 3a) and not in short duration (200 ms, Experiment 2b and 3b). Above findings were also based on the stringent criteria of subjective and objective unawareness.

The affective priming effect of facial expression is consistent with previous studies that showed processing of facial expression in different level such as perceptual and appraisal influence under CFS. For example, Adams et al. (2010) found that invisible facial expressions, such as angry, fearful, and happy under CFS can induce facial expressions aftereffect on subsequent target face (but see E. Yang, Hong, et al., 2010). Also, the affective and personality (such as trustworthy, attractive, and interpersonally warm) judgment of neutral face (Anderson, Siegel, White, & Barrett, 2012) and preference judgment of unrecognizable word (Almeida, Pajtas, Mahon, Nakayama, & Caramazza, 2013) can be affected by invisible facial expression under CFS (but see Faivre, Berthet, & Kouider, 2012). Consistent with these studies, the current finding takes a further step to reveal that emotional content of facial expression can directly affect the emotional judgment of words at semantic level under CFS.

In contrast to previous studies (Kang et al., 2011; Zimba & Blake, 1983), which showed that semantic meaning of word cannot be extracted under suppression state, the current study revealed priming effect of emotional word under CFS. This result extends the limit that not only low-level information, but high level information such as semantic meaning can also be processed unconsciously. The contrary results may be attributed to a critical factor: the duration of invisible prime. In these cases, the prime durations under suppression were shorter (280ms of Exp. 3 in Kang et al., 2011; 64~194 ms in Zimba & Blake, 1983) than the current setting (1000 ms in Experiment 2b and 3a). Concerning the pros and cons of long suppression time in CFS, although some suggest that longer suppression time produce stronger suppression and thus the priming effect preserves at shorter but not long duration (E. Yang et al., 2014), others emphasize that longer duration elicits stronger sensory stimulation (Barbot & Kouider, 2012). Our data was in favor of the latter account, since we found that the priming effect of word was dependent on the prime duration. That is, unlike the affective priming effect of facial expression which can be kept at short duration (200 ms, Exp. 2a), the emotional effect of word was not found at short duration (200 ms, Exp. 2b and 3b). Therefore, if the semantic processing of words requires longer duration, then the CFS paradigm provides a better opportunity to prolong the unconscious processing of words than binocular rivalry.

Regarding the different priming pattern between face and word, we excluded the possibilities of cognitive aftereffect (Exp. 2b and 3b) and the response bias of face target (Exp. 3) for the negative priming in negative word. Another possibility may be due to different processing between face and word on the interpretation of the emotional prime. For the face prime, the specific emotion such as fearful and happy face were used, which could generate corresponding evaluation of emotion. On the other hand, the meaning of same emotional word may have various connotations to different observers, though it was labeled by us as negative or positive valence. In contrast to the negative valence, which included different sub-types of valences, the positive valence may all refer to ”happy”. Therefore, if different connotations of negative emotional meanings were extracted, which may interrupt the emotional judgment of subsequent target, it may account for the negative impairment in our result from negative prime words. Further studies are needed to consider the diverse connotations that may have embedded in word stimuli to test this possibility.

In conclusion, comparing to the b-CFS paradigm, the advantage of priming-CFS can clearly reveal the attribution of emotional meaning under interocular suppression. Studies conducted in the last decade using the priming-CFS paradigm have shown that information sent through visual dorsal stream, such as motion (Faivre & Koch, 2014; Khuu, Chung, Lord, & Pearson, 2014), tools (Almeida et al., 2010; Almeida et al., 2008; but see Sakuraba, Sakai, Yamanaka, Yokosawa, & Hirayama, 2012), and numerical information (Bahrami et al., 2010; Sklar et al., 2012) can be processed unconsciously. Here, we demonstrated that the information passed through ventral stream, such as face and word, can also be processed unconsciously. Finally, since different priming effect was found between face and word, it is important to pay attention to the stimuli specific effect while different types of emotion-laden stimuli are used in future investigations.

General Discussion

Summary of Results

From the findings of three studies, we found that emotional content of facial expression can be extracted under CFS. In Study 1 and Study 2, we used direct measurement (i.e., breaking-CFS and physiology-CFS) with subjective criteria of unconscious processing, and found that meaningfulness (i.e., own-race familiarity and multisensory integration) can modulate unconscious processing of facial expression under CFS. In Study 3, we adopted indirect measurement (i.e., priming-CFS) with both subjective and objective criteria of unconscious processing, and found that congruency of invisible emotional information under CFS can influence emotional judgment of subsequent target. These results support that facial expression—which is supposedly processed along ventral pathway—can be processed unconsciously.

There are different levels of information embedded in a facial expression, and thus it is important to figure out which level of processing is involved when emotional content is extracted unconsciously. For example, mouth region is proposed to be important for recognizing happy face (Beaudry, Roy-Charland, Perron, Cormier, & Tapp, 2014; Calvo, Fernandez-Martin, & Nummenmaa, 2012; T. Stein & Sterzer, 2012). This would be due to the low-level property of parallel curvature between mouth and face contours that are irrelevant to emotion (T. Stein & Sterzer, 2012), or local salient features that represent smiling at the stage of perceptual processing (Beaudry et al., 2014; Calvo et al., 2012), or affective evaluation of positive valence (Calvo et al., 2012; Russell, 2003). Current studies respectively tease apart the contributions from low-level properties, local-salient features, and the affective evaluation of facial expression.

In Study 1, we found that own-race familiarity, but not orientation familiarity, can modulate the processing of facial expression under CFS. In the literature, orientation familiarity has been well adopted as a control condition under CFS; however, most studies do not find the modulation effect in facial expression (e.g., Gray et al., 2013; T. Stein & Sterzer, 2012; E. Yang et al., 2007), and this is further supported in the current study. On the other hand, we found that own-race familiarity can modulate the processing of facial expression under CFS, and this finding suggests that own-race familiarity is a reliable control to interrupt the meaningfulness of facial expression but keep low-level features relatively unaffected. More importantly, since there may not be so-called "familiar curvature" (for mouth region in happy face, T. Stein & Sterzer, 2012) or "familiar contrast" (for eye-white region in fearful face, Gray et al., 2013), this study excluded the explanation of low-level properties that contribute to our results.

In Study 2 we found that voice can facilitate face processing under CFS at different levels. In Experiment 1, we found that only human voice, but not beep sound, can facilitate unconscious processing of upright face. In Experiment 2, as revealed by the congruency effect between affective voice and facial expression, we found that this unconsciously voice-face binding can be processed to specific emotion. These results extended the finding of the modulation effect of meaningfulness in Study 1 from visual modality to multisensory integration.

This multisensory enhancement provides more convincing evidence for the conclusion that emotional content can be processed under CFS. In Study 1, although Asian faces and Caucasian faces were generated from the same database to minimize the difference in low-level properties, the processing of facial expression was still compared between different sets of faces (i.e., own-race face versus other-race face), which may involve different features among faces. In Study 2, the faces were the same between congruent and incongruent conditions, but the processing efficiency was modulated by the congruency of affective voice. Therefore, different suppression times within the same facial expression were resulted from the congruent emotional content of multisensory enhancement without covariant with any low-level properties. It may be argued that the affective voice was integrated with local-salient features, such as hearing the screaming sound or laughing sound then looking for the eye-white region of fearful face or mouth region of happy face, respectively. However, as mentioned above, the local-salient features (e.g., eye-white region in fearful face) are not equivalent to low-level properties (e.g., contrast): the former ones are assumed to be tightly connected with emotional recognition (Adolphs et al., 2005; Beaudry et al., 2014). Again, since curvature or contrast should not integrate with affective voice, this study also excluded the contribution of low-level properties to the results.

In Study 3, we used the priming paradigm under CFS and found that invisible emotional information, laden by face and word, affected emotional judgment of subsequent target. For the invisible face prime, prime with a long duration (1000ms) and a short duration (200ms) consistently revealed a positive priming effect. On the other hand, positive and negative word prime revealed positive priming and negative priming, respectively, and the word priming effects were only found in the long duration (1000ms) but not short duration (200ms) conditions. These results suggest that even though face and word have different priming effect and processing efficiency, their emotional content can nevertheless be extracted under CFS. Compared to Study1 and Study 2, affective-priming procedure directly revealed that affective evaluation of emotional information were extracted under CFS. Moreover, the findings in Study 3 are based on more stringent criteria to define unconscious processing, that is, both subjective and objective measurements.

Revisited the Processing Capacity under CFS

It has been long debated about the issue whether high-level information can be processed under interocular suppression. Leading studies from monkey electrophysiology (Sheinberg & Logothetis, 1997) and human *f*MRI (Tong et al., 1998) both showed attenuated neuronal response along the ventral pathway when the information was suppressed during binocular rivalry. However, recent studies by adopting CFS revealed very different pictures on this issue. While some suggested that facial expression can be processed under CFS (Adams et al., 2010; E. Yang et al., 2007), others argued that the supported results may be confounded with low-level properties (Gray et al., 2013; T. Stein & Sterzer, 2012) or partial consciousness (T. Stein & Sterzer, 2014; E. Yang, Hong, et al., 2010).

Inspired by these studies, in the current thesis we addressed this issue by adopting different approaches. In Study 1 and Study 2, unlike previous studies (e.g., Gray et al., 2013; E. Yang et al., 2007) using facial expression alone without finding different results between formal condition and control condition, we measured the modulation effect of meaningfulness on the processing of facial expression. In addition, convergent evidence between breaking-CFS and eye-movements supports that breaking-CFS is a reliable index of unconscious processing. In Study 3, we used both subjective and objective criteria to define unconscious processing, and using indirect measurement to test the affective priming effect. These approaches do not confound emotional information with low-level properties, and thus provide more convincing evidence for this issue.

Some brain areas, such as superior temporal sulcus (STS) and amygdala, may involve in the processing of facial expression under CFS, however, the exact brain-loci remains unclear for this processing due to the results seem highly dependent on individual difference. For example, Jiang and He (2006) compared fearful face, neutral face and scramble face under CFS with *f*MRI measurement, and found that fearful face led to higher activation in STS and bilateral amygdala than neutral face. However, Vizueta, Patrick, Jiang, Thomas, and He (2012) used similar settings as Jiang and He (2006), but the activation of STS and amygdala cannot differentiate fearful face and neutral face. Interestingly, they found that only participants with high negative affect revealed different activations of these two areas between fearful face and neutral face. More specially, participant with amygdala lesion can still process facial expression by adopting breaking-CFS (Tsuchiya et al., 2009). This finding suggests that amygdala may not be necessary for initial detection of facial expression under CFS. These results suggest that individual difference may modulate how neural system processes affective information under CFS. Here, we provide a solution for future studies (see below, the "The proposed gaze-contingent -CFS paradigm" section, p. 90).

Across three studies, we not only found that emotional content of facial expression can be processed under interocular suppression, but also found that facial expression can interact with other information at different levels. The own-race familiarity examined in Study 1 is more relevant to facial identity (Elfenbein & Ambady, 2002; Matsumoto, 1989). Although both are related to recognition of the same face, processing of facial expression and that of facial identity are distinct (Bruce & Young, 1986; Haxby et al., 2000). In Study 3, the priming effect found between facial expression and emotional word was based on different types of stimuli. Although both were processed within visual modality, face (fusiform face area; Kanwisher, McDermott, & Chun, 1997) and word (visual word form area; Cohen et al., 2000) are processed in distinctive brain areas. In Study 2, the affective voice we used triggered a kind of voice perception in auditory modality (Belin et al., 2004). Although they are naturally connected in emotional recognition, the congruency between facial expression and affective voice was cross-modalities

(Campanella & Belin, 2007; de Gelder & Vroomen, 2000). These findings suggest that information through the ventral pathway can not only be processed unconsciously but also cross-talked in widely spread brain areas.

Goodale (2007) recently proposed that ventral stream and dorsal steam may play different roles on conscious and unconscious processing, respectively. Ventral stream is relevant to conscious perception of object recognition, providing "off-line" processing of visual information, and not for unconscious processing such as automatic response. On the other hand, dorsal stream controls our on-line action, and most of our action can be accomplished without access to consciousness (Goodale, 2007). Extending this theory (Goodale, 2007), we conclude that although processing along ventral stream is necessary but it is not sufficient to generate conscious perception.

Our results also support neurobiological theory of consciousness such as the global neuronal workspace theory (GNWT; Dehaene et al., 2006; Dehaene & Naccache, 2001). This theory suggests that there are two types of processing, one is the local areas that are responsible for specific processing, and the other is more "global" workspace areas that can receive and integrate different information from local areas. Conscious perception only occurs with strong sensory input *and* focused top-down attention. In our case, although participants paid their attention on the tasks, the suppressed information under CFS might be too weak for conscious access. According to the GNWT, although the weak sensory input cannot be accessed, the feedforward processing can spread across different brain areas. Since information such as facial expression, familiarity, word and voice can be separately encoded within or across modalities, this explanation suggests a possibility that information can be integrated unconsciously (Z. Lin & He, 2009; Mudrik, Faivre, & Koch, 2014).

Implications on the Issue of Consciousness

Our findings revealed that information under CFS is a cumulative processing to break interocular suppression and to influence the processing of subsequent target. For example, Study 1 showed that unfamiliar race and unfamiliar orientation of face required longer breaking-CFS time than its counterpart. Study 2 showed that incongruent affective face-voice paring also required longer processing than congruent paring to access consciousness. Study 3 found that emotional word can only generate affective-priming effect at a long duration but not short duration. These studies all suggest that unfamiliar, unexpected, and complex information required more time to be processed unconsciously (see also Lo & Yeh, 2008) and accessed for consciousness. Standing on this point, unconscious processing is a prerequisite for conscious access. In other words, conscious processing has no executive function for information processing (Earl, 2014), and consciousness and unconsciousness differ quantitatively but not qualitatively. This speculation is based on current findings that facial expression can be processed and integrated with different information unconsciously (cf., Tononi & Koch, 2008; van Boxtel et al., 2010).

What is the function of consciousness, then? We believe that the function of consciousness is more relevant to flexible and voluntary responses to the relevant information in a task. For example, McCormick (1997) has probed how conscious states modulate the attention in a revised attentional orientating task (Posner, 1980). In his study, he manipulated a spatial cue having 85% probability that can predict a target presented on its opposite location. Therefore, participants needed to control their attention to the opposite location of cue to detect the target. He found that the participants can endogenously shift their attention to the opposite location of cue when the cue was consciously perceived, but attention was exogenously driven by cue location (even with only 15% probability) when the cue was unconsciously perceived. This result implies that one of the functions of consciousness is related to the switch of flexible response when information violates automatic response.

In the current thesis, the measurements of conscious and unconscious processing were within the same context of information processing. For example, in Study 3 we asked the participants to detect the prime (for excluding the subjective awareness) and make a forced-choice to the valence of prime (for excluding objective awareness) and used the affective-priming procedure to access unconscious processing of prime. All of the measurements were relevant to the similar response to the prime, it is thus not necessary using flexible response in the task. Although our results overturned the idea that the function of consciousness is to integrate and discriminate information (Balduzzi & Tononi, 2008; Tononi, 2004; van Boxtel et al., 2010), we only found quantitative difference between conscious and unconscious processing. Further studies are needed to adopt a task requiring flexible response to reveal qualitative difference between conscious and unconscious processing (Schmidt & Vorberg, 2006).

The Proposed Gaze-Contingent CFS Paradigm

In this thesis, we adopted three kinds of measurements under CFS, and found that emotional content of facial expression can be processed under CFS. While CFS is a well adopted paradigm to probe unconscious processing in recent years, it is still hotly debated about what kind of index is a suitable measurement (see Faivre et al., 2014; Gayet et al., 2014; T. Stein & Sterzer, 2014; E. Yang et al., 2014 for discussions). Based on our findings and the advantages of different kinds of measurement, we proposed a new approach called the *gaze-contingent CFS* paradigm for future studies. Like the gaze-contingent crowding paradigm (Faivre & Kouider, 2011b) which monitors participant's gaze to replace crowed information in the spatial domain, the proposed gaze-contingent CFS paradigm monitors participants' gaze to replace suppressed information in the temporal domain. The idea of the gaze-contingent CFS paradigm is based on the observation of the eye-movement results in Study 2, according to which the first-saccade latency can predict the subsequent suppression time of breaking-CFS before 1000 ms. Therefore, first-saccade latency provides a reliable index of unconscious processing under CFS. Gaze-contingent CFS adopts first-saccade latency as an index to define a suitable duration for individual difference and different stimuli. That is, when participants move their eyes to a stimulus location for a period of time, assuming the information has been processed (as shown in the results of Study 2), then the continuously flashed masks stops automatically. After that, the subjective and/or objective measurement can be used as the criteria to define unconscious processing.

Finding a suitable suppression time is important not only for successful suppression across different individuals, but also for the degree of efficiency that information is processed. For example, different participants lead to large variance of breaking-CFS times (see the result of Study 2, p.52), and different breaking-CFS time also infers that processing efficiencies are varied across different participants. Moreover, short suppression time may induce insufficient processing of information in some cases (e.g., word prime in Study 3), but long duration may induce over-stimulation in other cases (e.g., Barbot & Kouider, 2012). Therefore, the optimal suppression time for information processing not only varies by individual difference but also by the types of information. This interaction becomes very complicated for finding a suitable duration when probing different levels of processing across different participants. As mentioned in the section of "measurement under CFS" (p.13), the fixed duration adopted by aftereffect-CFS cannot fulfill this requirement.

The proposed gaze-contingent CFS paradigm in this thesis combines several advantages from different measurements. On the one hand, it provides a flexible suppression (also processing) time like breaking-CFS. On the other hand, eye-movements can be a measurement of unconscious processing and they can also be an index for the optimal suppression duration under CFS. Moreover, after the gaze-contingent CFS manipulation, indirect measurement like priming-CFS can also be tested. Finally, this paradigm can adopt stringent criteria such as objective measurement to define unconscious processing.

Conclusions

The current thesis provides new evidence for the issue of the extent that information can be processed under interocular suppression. Findings across three studies support that information along the ventral stream can be processed under CFS. By manipulating the meaningfulness of facial expression in Study 1 and Study 2, we excluded the explanation of low-level properties (e.g., contrast, curvature) that may influence the unconscious processing of facial expression (Gray et al., 2013; T. Stein & Sterzer, 2012). Note that although the local-salient features representing emotional information may play a partial role to render facial expression that being accessed into consciousness in both studies, previous studies also suggest that affective evaluation can be automatically extracted from facial expression (Calvo et al., 2012; Carroll & Young, 2005). This idea was further supported in Study 3, where results from the affective-priming procedure directly showed that affective evaluation of emotional information were extracted under CFS and influenced the emotional judgment of subsequent target. Our findings overturn the common believe that consciousness is needed for information integration (Balduzzi & Tononi, 2008; Tononi, 2012; Tononi & Koch, 2008) and suggest that some information can be processed and integrated unconsciously to a certain extent. Finally, different measurements across the three studies reveal coherent findings. These observations enlighten us to propose a new gaze-contingent CFS paradigm to measure unconscious processing under CFS. Combining several advantages from different measurements, the gaze-contingent CFS paradigm may provide a powerful tool to measure unconscious processing in the future.

Acknowledgments

This work was supported by Taiwan's Ministry of Science and Technology under Grant MOST101-2410-H-002-083-MY3 to Su-Ling Yeh and MOST103-2420-H-002-027-DR to Yung-Hao Yang.

References

Adams, W. J., Gray, K. L., Garner, M., & Graf, E. W. (2010). High-level face adaptation without awareness. *Psychological Science, 21*(2), 205-210.

Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology, 12*(2), 169-177.

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature, 433*(7021), 68-72.

Alais, D., & Blake, R. (2005). *Binocular rivalry*. Cambridge, MA: MIT Press.

Almeida, J., Mahon, B. Z., & Caramazza, A. (2010). The role of the dorsal visual processing stream in tool identification. *Psychological Science, 21*(6), 772-778.

Almeida, J., Mahon, B. Z., Nakayama, K., & Caramazza, A. (2008). Unconscious processing dissociates along categorical lines. *Proceedings of the National Academy of Sciences of the United States of America, 105*(39), 15214-15218.

Almeida, J., Pajtas, P. E., Mahon, B. Z., Nakayama, K., & Caramazza, A. (2013). Affect of the unconscious: visually suppressed angry faces modulate our decisions. *Cognitive, Affective & Behavioral Neuroscience, 13*(1), 94-101.

Anderson, E., Siegel, E., White, D., & Barrett, L. F. (2012). Out of sight but not out of mind: unseen affective faces influence evaluations and social impressions. *Emotion, 12*(6), 1210-1221.

Baars, B. J. (2005). Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. *Progress in Brain Research, 150*, 45-53.

Bahrami, B., Vetter, P., Spolaore, E., Pagano, S., Butterworth, B., & Rees, G. (2010). Unconscious numerical priming despite interocular suppression. *Psychological Science, 21*(2), 224-233.

Balduzzi, D., & Tononi, G. (2008). Integrated information in discrete dynamical systems: motivation and theoretical framework. *PLOS Computational Biology, 4*(6), e1000091.

Barbot, A., & Kouider, S. (2012). Longer is not better: nonconscious overstimulation reverses priming influences under interocular suppression. *Attention, perception & psychophysics, 74*(1), 174-184.

Baudouin, J. Y., Sansone, S., & Tiberghien, G. (2000). Recognizing expression from familiar and unfamiliar faces. *Pragmatics & Cognition, 8*(1), 123-146.

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.

Belin, P., Fecteau, S., & Bedard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends in Cognitive Sciences, 8*(3), 129-135.

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.

Biehl, M., Matsumoto, D., Ekman, P., Hearn, V., Heider, K., Kudoh, T., et al. (1997). Matsumoto and Ekman's Japanese and Caucasian facial expressions of emotion (JACFEE): reliability data and cross-national differences. . *Journal of Cross-Cultural Psychology*(21), 3–21.

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS One, 3*(8), e3022.

Blake, R. (1988). Dichoptic reading: the role of meaning in binocular rivalry. *Perception & Psychophysics, 44*(2), 133-141.

Blake, R., & Fox, R. (1974). Adaptation to invisible gratings and the site of binocular rivalry suppression. *Nature, 249*(456), 488-490.

Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience, 3*(1), 13-21.

Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences, 9*(2), 46-52.

Bonneh, Y. S., Cooperman, A., & Sagi, D. (2001). Motion-induced blindness in normal observers. *Nature, 411*(6839), 798-801.

Bridge, H., Thomas, O., Jbabdi, S., & Cowey, A. (2008). Changes in connectivity after visual cortical brain damage underlie altered visual function. *Brain, 131*(Pt 6), 1433-1444.

Brigard, F. D., & Prinz, J. (2010). Attention and consciousness. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(1), 51-59.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77 (Pt 3)*, 305-327.

Calvo, M. G., & Beltran, D. (2013). Recognition advantage of happy faces: tracing the neurocognitive processes. *Neuropsychologia, 51*(11), 2051-2061.

Calvo, M. G., Fernandez-Martin, A., & Nummenmaa, L. (2012). Perceptual, categorical, and affective processing of ambiguous smiling facial expressions. *Cognition, 125*(3), 373-393.

Calvo, M. G., & Lundqvist, D. (2008). Facial expressions of emotion (KDEF): identification under different display-duration conditions. *Behavior Research Methods, 40*(1), 109-115.

Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences, 11*(12), 535-543.

Carroll, N. C., & Young, A. W. (2005). Priming of emotion recognition. *The Quarterly journal of experimental psychology. A, Human Experimental Psychology, 58*(7), 1173-1197.

Chalmers, D. J. (1998). On the Search for the Neural Correlate of Consciousness. Paper presented at the Toward a Science of Consciousness II: The Second Tucson Discussions and Debates.

Chen, L. F., & Yen, Y. S. (2007). Taiwanese facial expression image database. Taipei, Taiwan.: Brain Mapping Laboratory, Institute of Brain Science, National Yang-Ming University.

Chen, Y. C., & Yeh, S. L. (2008). Visual events modulated by sound in repetition blindness. *Psychonomic Bulletin & Review, 15*(2), 404-408.

Chen, Y. C., & Yeh, S. L. (2012). Look into my eyes and I will see you: unconscious processing of human gaze. *Consciousness and Cognition, 21*(4), 1703-1710.

Chen, Y. C., Yeh, S. L., & Spence, C. (2011). Crossmodal constraints on human perceptual awareness: auditory semantic modulation of binocular rivalry. *Frontiers in Psychology, 2*, 212.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., et al. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain, 123 (2)*, 291-307.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82*, 407-428.

Costello, P., Jiang, Y., Baartman, B., McGlennen, K., & He, S. (2009). Semantic and subword priming during binocular suppression. *Consciousness and Cognition, 18*(2), 375-382.

de Gelder, B., Morris, J. S., & Dolan, R. J. (2005). Unconscious fear influences emotional awareness of faces and voices. *Proceedings of the National Academy of Sciences of the United States of America, 102*(51), 18682-18687.

de Gelder, B., Pourtois, G., Vroomen, J., & Bachoud-Levi, A. C. (2000). Covert processing of faces in prosopagnosia is restricted to facial expressions: Evidence from cross-modal bias. . *Brain Cognition, 44*(3), 425-444.

de Gelder, B., Pourtois, G., & Weiskrantz, L. (2002). Fear recognition in the voice is modulated by unconsciously recognized facial expressions but not by unconsciously recognized affective pictures. *Proceedings of the National Academy of Sciences of the United States of America, 99*(6), 4121-4126.

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

de Gelder, B., Vroomen, J., & Bertelson, P. (1998). Upright but not inverted faces modify the perception of emotion in the voice. *Current Psychology of Cognition, 17*, 1021-1031.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences, 10*(5), 204-211.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition, 79*(1-2), 1-37.

Del Cul, A., Dehaene, S., & Leboyer, M. (2006). Preserved subliminal processing and impaired conscious access in schizophrenia. *Archives of General Psychiatry, 63*(12), 1313-1323.

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 of the United States of America, 98*(17), 10006-10010.

Earl, B. (2014). The biological function of consciousness. *Front Psychol, 5*, 697.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*(2), 124-129.

Ekman, P., & Friesen, W. V. (1976). Pictures of facial affect. Palo Alto, CA: Consulting Psychologist Press. In P. Alto (Ed.), *CA*: Consulting Psychologists.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: a meta-analysis. *Psychological Bulletin, 128*(2), 203-235.

Elfenbein, H. A., & Ambady, N. (2003). When familiarity breeds accuracy: cultural exposure and facial emotion recognition. *Journal of Personality and Social Psychology, 85*(2), 276-290.

Fahle, M. (1982a). Binocular rivalry: suppression depends on orientation and spatial frequency. *Vision Research, 22*(7), 787-800.

Fahle, M. (1982b). Cooperation between different spatial frequencies in binocular rivalry. *Biological Cybernetics, 44*(1), 27-29.

Faivre, N., Berthet, V., & Kouider, S. (2012). Nonconscious influences from emotional faces: a comparison of visual crowding, masking, and continuous flash suppression. *Frontiers in Psychology, 3*, 129.

Faivre, N., Berthet, V., & Kouider, S. (2014). Sustained invisibility through crowding and continuous flash suppression: a comparative review. *Frontiers in Psychology, 5*, 475.

Faivre, N., & Koch, C. (2014). Temporal structure coding with and without awareness. *Cognition, 131*(3), 404-414.

Faivre, N., & Kouider, S. (2011a). Increased sensory evidence reverses nonconscious priming during crowding. *Journal of Vision, 11*(13).

Faivre, N., & Kouider, S. (2011b). Multi-feature objects elicit nonconscious priming despite crowding. *Journal of Vision, 11*(3).

Fang, F., & He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nature Neuroscience, 8*(10), 1380-1385.

Fendrich, R., Wessinger, C.M. and Gazzaniga, M.S. (2001). Speculations on the neural basis of islands of blindsight. *Progress in Brain Research, 134*, 353-366.

Gayet, S., Paffen, C. L., & Van der Stigchel, S. (2013). Information matching the content of visual working memory is prioritized for conscious access. *Psychological Science, 24*(12), 2472-2480.

Gayet, S., Van der Stigchel, S., & Paffen, C. L. (2014). Breaking continuous flash suppression: competing for consciousness on the pre-semantic battlefield. *Frontiers in Psychology, 5*, 460.

Gibson, J. J., & Radner, M. (1937). Adaptation, aftereffect and contrast in the perception of tilted lines. *Journal of Experimental Psychology, 20*, 453-467.

Gobbini, M. I., Gors, J. D., Halchenko, Y. O., Rogers, C., Guntupalli, J. S., Hughes, H., et al. (2013). Prioritized Detection of Personally Familiar Faces. *PLoS One, 8*(6), e66620.

Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia, 45*(1), 32-41.

Goodale, M. A. (2007). Duplex vision: Separate cortical pathways for conscious perception and the control of action. In M. Velmans & S. Schneider (Eds.), *The Blackwell Companion to Consciousness* (pp. 616-627). Oxford: Blackwell.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15*(1), 20-25.

Gray, K. L., Adams, W. J., Hedger, N., Newton, K. E., & Garner, M. (2013). Faces and awareness: Low-level, not emotional factors determine perceptual dominance. *Emotion, 13*(3), 537-544.

Hare, T. A., Tottenham, N., Davidson, M. C., Glover, G. H., & Casey, B. J. (2005). Contributions of amygdala and striatal activity in emotion regulation. *Biological Psychiatry, 57*(6), 624-632.

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.

Hsiao, J. Y., Chen, Y. C., Spence, C., & Yeh, S. L. (2012). Assessing the effects of audiovisual semantic congruency on the perception of a bistable figure. *Consciousness and Cognition, 21*(2), 775-787.

Hsieh, P. J., Colas, J. T., & Kanwisher, N. (2011). Pop-out without awareness: unseen feature singletons capture attention only when top-down attention is available. *Psychological Science, 22*(9), 1220-1226.

Hsu, L. C., Yeh, S. L., & Kramer, P. (2004). Linking motion-induced blindness to perceptual filling-in. *Vision Research, 44*(24), 2857-2866.

Hsu, L. C., Yeh, S. L., & Kramer, P. (2006). A common mechanism for perceptual filling-in and motion-induced blindness. *Vision Research, 46*(12), 1973-1981.

Huber, D. E. (2008). Immediate priming and cognitive aftereffects. *Journal of Experimental Psychology. General, 137*(2), 324-347.

Huber, D. E. (2015). Using continual flash suppression to investigate cognitive aftereffects. *Consciousness and Cognition, 35*, 30-32.

Huber, D. E., & O'Reilly, R. C. (2003). Persistence and accommodation in short-term priming and other perceptual paradigms: temporal segregation through synaptic depression. *Cognitive Science, 27*(3), 403-430.

Izard, C. E. (1994). Innate and universal facial expressions: evidence from developmental and cross-cultural research. *Psychological Bulletin, 115*(2), 288-299.

Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology, 19*(18), 1543-1548.

Jiang, Y., Costello, P., Fang, F., Huang, M., & He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences of the United States of America, 103*(45), 17048-17052.

Jiang, Y., Costello, P., & He, S. (2007). Processing of invisible stimuli: advantage of upright faces and recognizable words in overcoming interocular suppression. *Psychological Science, 18*(4), 349-355.

Jiang, Y., Fang, F., Huang, M., & He, S. (2005). Invisible interesting pictures can attract spatial attention. *Journal of Vision, 5*(8), 511.

Jiang, Y., & He, S. (2006). Cortical responses to invisible faces: dissociating subsystems for facial-information processing. *Current Biology, 16*(20), 2023-2029.

Kanai, R., Walsh, V., & Tseng, C. H. (2010). Subjective discriminability of invisibility: a framework for distinguishing perceptual and attentional failures of awareness. *Consciousness and Cognition, 19*(4), 1045-1057.

Kang, M. S., Blake, R., & Woodman, G. F. (2011). Semantic analysis does not occur in the absence of awareness induced by interocular suppression. *The Journal of Neuroscience, 31*(38), 13535-13545.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience, 17*(11), 4302-4311.

Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition, 68*(1), B1-11.

Kao, C. H., Chen, D. Y., & Chen, C. C. (2010). The inversion effect in visual word form processing. *Cortex, 46*(2), 217-230.

Kaunitz, L., Fracasso, A., Lingnau, A., & Melcher, D. (2013). Non-conscious processing of motion coherence can boost conscious access. *PLoS One, 8*(4), e60787.

Khuu, S. K., Chung, C. Y., Lord, S., & Pearson, J. (2014). Unconscious local motion alters global image speed. *PLoS One, 9*(12), e112804.

Kim, C. Y., & Blake, R. (2005). Psychophysical magic: rendering the visible 'invisible'. *Trends in Cognitive Sciences, 9*(8), 381-388.

Koch, C., & Crick, F. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences, 2*, 263-275.

Koch, C., & Crick, F. (2001). Neural Basis of Consciousness. In N. J. Smelser & P. B. Baltes (Eds.), International Encyclopedia of the Social & Behavioral Sciences (pp. 2600-2604): Elsevier Science Ltd.

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in Cognitive Sciences, 11*(1), 16-22.

Koch, C., & Tsuchiya, N. (2012). Attention and consciousness: related yet different. *Trends in cognitive sciences, 16*(2), 103-105.

Kouider, S. (2009). Neurobiological Theories of Consciousness. In W. Banks (Ed.), *Encyclopedia of Consciousness* (pp. 87-100): Elsevier.

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.

Kreifelts, B., Ethofer, T., Shiozawa, T., Grodd, W., & Wildgruber, D. (2009). Cerebral representation of non-verbal emotional perception: fMRI reveals audiovisual integration area between voice- and face-sensitive regions in the superior temporal sulcus. *Neuropsychologia, 47*(14), 3059-3066.

Kunimoto, C., Miller, J., & Pashler, H. (2001). Confidence and accuracy of near-threshold discrimination responses. *Consciousness and Cognition, 10*(3), 294-340.

Lamme, V. A. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences, 7*(1), 12-18.

Lamme, V. A. (2004). Separate neural definitions of visual consciousness and visual attention; a case for phenomenal awareness. *Neural Networks, 17*(5-6), 861-872.

Laureys, S. (2005). The neural correlate of (un)awareness: lessons from the vegetative state. *Trends in Cognitive Sciences, 9*(12), 556-559.

Lehmkuhle, S. W., & Fox, R. (1975). Effect of binocular rivalry suppression on the motion aftereffect. *Vision Research, 15*(7), 855-859.

Levi, D. M. (2008). Crowding--an essential bottleneck for object recognition: a mini-review. *Vision Research, 48*(5), 635-654.

Li, J. L., & Yeh, S. L. (2007). New objects do not capture attention without a top-down setting: Evidence from an inattentional blindness task. *Visual Cognition, 15*, 661-684.

Lin, S. Y., & Yeh, S. L. (2015). Unconscious grouping of Chinese words: Evidence from object-based attention. *Language and Linguistics, 16*(4), 517-533.

Lin, Z., & He, S. (2009). Seeing the invisible: the scope and limits of unconscious processing in binocular rivalry. *Progress in Neurobiology, 87*(4), 195-211.

Lo, S. Y., & Yeh, S. L. (2008). Dissociation of processing time and awareness by the inattentional blindness paradigm. *Consciousness and Cognition, 17*(4), 1169-1180.

Mack, A., & Rock, I. (1998). *Inattentional blindness*. MA: MIT Press.: Cambridge.

Marcel, A. J. (1983). Conscious and unconscious perception: experiments on visual masking and word recognition. *Cognitive Psychology, 15*(2), 197-237.

Matsumoto, D. (1989). Cultural influences on the perception of emotion. *Journal of Cross-Cultural Psychology, 20*, 92-105.

McCormick, P. A. (1997). Orienting attention without awareness. *Journal of Experimental Psychology. Human Perception and Performance, 23*(1), 168-180.

McNamara, T. P. (1992). Theories of priming: I. Associative distance and lag. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(6), 1173-1190.

Merikle, P. M., & Daneman, M. (2000). Conscious vs. unconscious perception. In M. S. Gazzaniga (Ed.), *The New Cognitive Neurosciences: 2nd Edition* (pp. 1295-1303): Mit Press.

Merikle, P. M., Smilek, D., & Eastwood, J. D. (2001). Perception without awareness: perspectives from cognitive psychology. *Cognition, 79*(1-2), 115-134.

Milner, A. D., & Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia, 46*(3), 774-785.

Mogensen, J. (2011). Reorganization of the injured brain: implications for studies of the neural substrate of cognition. *Frontiers in Psychology, 2*, 7.

Mudrik, L., Breska, A., Lamy, D., & Deouell, L. Y. (2011). Integration without awareness: expanding the limits of unconscious processing. *Psychological Science, 22*(6), 764-770.

Mudrik, L., Faivre, N., & Koch, C. (2014). Information integration without awareness. *Trends in Cognitive Sciences, 18*(9), 488-496.

Mueller, T. J., & Blake, R. (1989). A fresh look at the temporal dynamics of binocular rivalry. *Biological Cybernetics, 61*(3), 223-232.

Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review, 4*, 435-450.

Overgaard, M. (2011). Visual experience and blindsight: a methodological review. *Experimental Brain Research, 209*(4), 473-479.

Polonsky, A., Blake, R., Braun, J., & Heeger, D. J. (2000). Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nature Neuroscience, 3*(11), 1153-1159.

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology, 32*(1), 3-25.

Pourtois, G., de Gelder, B., Bol, A., & Crommelinck, M. (2005). Perception of facial expressions and voices and of their combination in the human brain. *Cortex, 41*(1), 49-59.

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.

Prioli, S. C., & Kahan, T. A. (2015). Identifying words that emerge into consciousness: Effects of word valence and unconscious previewing. *Consciousness and Cognition, 35*, 88-97.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: an attentional blink? *Journal of Experimental Psychology. Human perception and performance, 18*(3), 849-860.

Rees, G., Kreiman, G., & Koch, C. (2002). Neural correlates of consciousness in humans. *Nature Reviews. Neuroscience, 3*(4), 261-270.

Reingold, E. M., & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception & Psychophysics, 44*(6), 563-575.

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition, 47*(1), 25-57.

Roseboom, W., & Arnold, D. H. (2011). Learning to reach for 'invisible' visual input. *Current Biology, 21*(13), R493-494.

Rothkirch, M., Stein, T., Sekutowicz, M., & Sterzer, P. (2012). A direct oculomotor correlate of unconscious visual processing. *Current Biology, 22*(13), R514-515.

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

Sakuraba, S., Sakai, S., Yamanaka, M., Yokosawa, K., & Hirayama, K. (2012). Does the human dorsal stream really process a category for tools? *The Journal of Neuroscience, 32*(11), 3949-3953.

Salomon, R., Lim, M., Herbelin, B., Hesselmann, G., & Blanke, O. (2013). Posing for awareness: proprioception modulates access to visual consciousness in a continuous flash suppression task. *Journal of vision, 13*(7), 2.

Sato, W., Kochiyama, T., & Yoshikawa, S. (2011). The inversion effect for neutral and emotional facial expressions on amygdala activity. *Brain Research, 1378*, 84-90.

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice, 9*(3), 235-248.

Schmidt, T., & Vorberg, D. (2006). Criteria for unconscious cognition: three types of dissociation. *Perception & Psychophysics, 68*(3), 489-504.

Sengpiel, F., Blakemore, C., & Harrad, R. (1995). Interocular suppression in the primary visual cortex: a possible neural basis of binocular rivalry. *Vision Research, 35*(2), 179-195.

Sengpiel, F., & Vorobyov, V. (2005). Intracortical origins of interocular suppression in the visual cortex. *The Journal of Neuroscience, 25*(27), 6394-6400.

Sheinberg, D. L., & Logothetis, N. K. (1997). The role of temporal cortical areas in perceptual organization. *Proceedings of the National Academy of Sciences of the United States of America, 94*(7), 3408-3413.

Sklar, A. Y., Levy, N., Goldstein, A., Mandel, R., Maril, A., & Hassin, R. R. (2012). Reading and doing arithmetic nonconsciously. *Proceedings of the National Academy of Sciences of the United States of America, 109*(48), 19614-19619.

Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science, 16*(3), 184-189.

Stein, B. E., Nancy, L., Wilkinson, L. K., & Price, D. D. (1996). Enhancement of perceived visual intensity by auditory stimuli: A psychophysical analysis. *Journal of Cognitive Neuroscience, 8*(6), 497-506.

Stein, T., End, A., & Sterzer, P. (2014). Own-race and own-age biases facilitate visual awareness of faces under interocular suppression. *Frontiers in Human Neuroscience, 8*, 582.

Stein, T., Hebart, M. N., & Sterzer, P. (2011). Breaking Continuous Flash Suppression: A New Measure of Unconscious Processing during Interocular Suppression? *Frontiers in Human Neuroscience, 5*, 167.

Stein, T., Senju, A., Peelen, M. V., & Sterzer, P. (2011). Eye contact facilitates awareness of faces during interocular suppression. *Cognition, 119*(2), 307-311.

Stein, T., & Sterzer, P. (2012). Not just another face in the crowd: Detecting emotional schematic faces during continuous flash suppression. *Emotion, 12*(5), 988-996.

Stein, T., & Sterzer, P. (2014). Unconscious processing under interocular suppression: getting the right measure. *Frontiers in Psychology, 5*, 387.

Stein, T., Sterzer, P., & Peelen, M. V. (2012). Privileged detection of conspecifics: evidence from inversion effects during continuous flash suppression. *Cognition, 125*(1), 64-79.

Sterzer, P., Hilgenfeldt, T., Freudenberg, P., Bermpohl, F., & Adli, M. (2011). Access of emotional information to visual awareness in patients with major depressive disorder. *Psychological Medicine, 41*(8), 1615-1624.

Tamietto, M., & de Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews. Neuroscience, 11*(10), 697-709.

Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron, 21*(4), 753-759.

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience, 5*(1), 42.

Tononi, G. (2012). Integrated information theory of consciousness: an updated account. *Archives Italiennes de Biologie, 150*(4), 293-329.

Tononi, G., & Koch, C. (2008). The neural correlates of consciousness: an update. *Annals of the New York Academy of Sciences, 1124*, 239-261.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., et al. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Research, 168*(3), 242-249.

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience, 8*(8), 1096-1101.

Tsuchiya, N., & Koch, C. (2008). The relationship between consciousness and attention. In S. Laureys & G. Tononi (Eds.), *The neurology of consciousness: Cognitive Neuroscience and Neuropathology* (pp. 63-77). London, UK: Academic Press.

Tsuchiya, N., Koch, C., Gilroy, L. A., & Blake, R. (2006). Depth of interocular suppression associated with continuous flash suppression, flash suppression, and binocular rivalry. *Journal of Vision, 6*(10), 1068-1078.

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.

Valentine, T. (1988). Upside-down faces: a review of the effect of inversion upon face recognition. *British Journal of Psychology, 79 (4)*, 471-491.

van Boxtel, J. J., Tsuchiya, N., & Koch, C. (2010). Consciousness and attention: on sufficiency and necessity. *Frontiers in Psychology, 1*, 217.

Van den Bussche, E., Hughes, G., Humbeeck, N. V., & Reynvoet, B. (2010). The relation between consciousness and attention: an empirical study using the priming paradigm. [Comparative Study]. *Consciousness and cognition, 19*(1), 86-97.

Vizueta, N., Patrick, C. J., Jiang, Y., Thomas, K. M., & He, S. (2012). Dispositional fear, negative affectivity, and neuroimaging response to visually suppressed emotional faces. *Neuroimage, 59*(1), 761-771.

Wade, N. J., & Wenderoth, P. (1978). The influence of colour and contour rivalry on the magnitude of the tilt after-effect. *Vision Research, 18*(7), 827-835.

Wang, L., Weng, X., & He, S. (2012). Perceptual grouping without awareness: superiority of Kanizsa triangle in breaking interocular suppression. *PLoS One, 7*(6), e40106.

Watanabe, M., Cheng, K., Murayama, Y., Ueno, K., Asamizuya, T., Tanaka, K., et al. (2011). Attention but not awareness modulates the BOLD signal in the human V1 during binocular suppression. *Science, 334*(11), 829-831.

Weiskrantz, L. (1996). Blindsight revisited. *Current Opinion in Neurobiology, 6*(2), 215-220.

Weiskrantz, L. (2000). Blindsight: implications for the conscious experience of emotion. In L. R.D. & N. L. (Eds.), *Cognitive Neuroscience of Emotion.* (pp. 277-295). New York: Oxford University Press.

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

Wild-Wall, N., Dimigen, O., & Sommer, W. (2008). Interaction of facial expressions and familiarity: ERP evidence. *Biological Psychology, 77*(2), 138-149.

Yang, E., & Blake, R. (2012). Deconstructing continuous flash suppression. *Journal of Vision, 12*(3), 8.

Yang, E., Blake, R., & McDonald, J. E., 2nd. (2010). A new interocular suppression technique for measuring sensory eye dominance. *Investigative Ophthalmology & Visual Science, 51*(1), 588-593.

Yang, E., Brascamp, J., Kang, M. S., & Blake, R. (2014). On the use of continuous flash suppression for the study of visual processing outside of awareness. *Frontiers in Psychology, 5*, 724.

Yang, E., Hong, S. W., & Blake, R. (2010). Adaptation aftereffects to facial expressions suppressed from visual awareness. *Journal of Vision, 10*(12), 24.

Yang, E., Zald, D. H., & Blake, R. (2007). Fearful expressions gain preferential access to awareness during continuous flash suppression. *Emotion, 7*(4), 882-886.

Yang, Q., Bucci, M. P., & Kapoula, Z. (2002). The latency of saccades, vergence, and combined eye movements in children and in adults. *Investigative Ophthalmology & Visual Science, 43*(9), 2939-2949.

Yang, Y.-H., & Yeh, S.-L. (2014). Unmasking the dichoptic mask by sound: spatial congruency matters. *Experimental Brain Research*, *232*(4), 1109-16.

Yang, Y. H., & Yeh, S.-L. (2011). Accessing the meaning of invisible words. *Consciousness and Cognition, 20*(2), 223-233.

Yeh, S. L., He, S., & Cavanagh, P. (2012). Semantic priming from crowded words. *Psychological Science, 23*(6), 608-616.

Zhou, G., Zhang, L., Liu, J., Yang, J., & Qu, Z. (2010). Specificity of face processing without awareness. *Consciousness and Cognition, 19*(1), 408-412.

Zimba, L. D., & Blake, R. (1983). Binocular rivalry and semantic processing: Out of sight, out of mind. *Journal of experimental psychology. Human perception and performance, 9*(5), 807-815.