The Dual Role of Culture on Signalling and Receiving Dynamic Facial Expressions

Chaona Chen

Submitted in fulfilment of the requirements for the

Degree of Ph.D.

School of Psychology

College of Science and Engineering

University of Glasgow

May 2017

Abstract

Human survival critically relies on communicating a broad set of social messages including physical states and mental states. The prerequisite for any successful social communication is the shared knowledge between the sender and the receiver about what and how a specific social signal is used. To communicate the broad set of social messages in daily life, human beings have developed complex facial movement patterns as one of the most important and powerful social signals. With increasing globalization, cross-cultural interactions are fast becoming integral to modern life, which presents increasing pressure for cross-cultural communication. Specifically, a broad set of facial expressions including conversational facial expressions is critical for clear communication because they guide the flow of social exchanges. Yet, our knowledge of such facial movement patterns is relatively limited in terms of their functions in different cultural context – for example, whether these important everyday facial expressions are understood across cultures or cause cross-cultural confusions. In this thesis, I explored how facial movement patterns are used in Western and East Asian culture to communicate a broad set of social messages including physical states and mental states. Specifically, I objectively characterized the structure of dynamic facial expression patterns using 4D computer graphics and a data-driven social psychophysics method. I then examined the role of culture in signaling and receiving facial expressions using the signal detection theory and Mutual Information analysis. Together, my results reveal for the first time how specific facial movement patterns are used to communicate a broad set of social messages in Western and East Asian culture and how culture shapes the signalling and perception of such facial expressions in cross-cultural communication. Finally, I discussed the implication of my results in the field of psychology, computer science and social robotics, with links to my future work on developing a mathematical model of face social signalling and transfer this knowledge to socially and culturally sensitive conversational agents.

Table of Contents

List of Tables

List of Figures

List of Supplemental Material

Acknowledgement

I still remember the first day of my Ph.D. when I came to my office room on the very top of a hill in the Glasgow city, feeling proud and excited about my new journey. The magnificent view of landscape from the window seems never changed over these years, neither my enthusiasm of being a researcher.

Research is certainly not an easy journey and no one can achieve the success alone. I would like to take this opportunity to thank a number of people for their support throughout my Ph.D. First, I would like to thank my supervisors, Dr. Rachael E. Jack and Prof. Philippe G. Schyns. Dr. Rachael E. Jack brings tremendous and exemplary dedication and honesty to everything she does, and this extends to her students. I could not make so many achievements over these years without her endless patience and guidance. I would like to thank Professor Philippe G. Schyns for his invaluable advice regarding data interpretation and guidance for my research career. I would like to thank Dr. Oliver Garrod for introducing me to a truly amazing field of computer graphics; Dr. Marry Ellen Foster and Dr. Amol Deshmukh for their valuable insights and suggestions towards my further work on social robots and conversational agents; Prof. Simon Garrod, Dr. Dale Barr, Dr. Christoph Scheepers, Dr. Guillaume A. Rousselet and Prof. Christoph Kayser for their invaluable suggestions and continued support during my final year. A special thank to Prof. James A. Russell and Prof. Lawrence W. Barsalou for have thoroughly examined my work via the informative and enjoyable discussion during my viva.

I have also been fortunate to collaborate with Professor Hongmei Yan in the University of Electronic Science and Technology in China, Professor José-Miguel Fernández-Dols, Dr. Carlos Crivelli in Universidad Autónoma de Madrid in Spain and Professor Daniel S. Messinger in University of Miami in the United States. I would like to thank them for their creative ideas and enthusiastic discussion throughout. Our collaborations are both fruitful and fun. I would like to thank my parents and my dearest friends, Dr. Miao Huang, Danielle Morrison, Dr. Ruolan Ouyang, Jia Xie, Dr. Liyu Cao for their continued encouragement and support on building my mental strength. Finally, I would like to thank both the Chinese Scholarship Council (201306270029) and the University of Glasgow for providing financial support for my Ph.D. research.

Every end is another beginning. My research career just started from here.

Author’s Declaration

This thesis represents the original work of Chaona Chen except where explicit reference is made to the contribution of others. I declare that this thesis is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

1 General Introduction

Man is by nature a social animal; an individual who is unsocial naturally and not accidentally is either beneath our notice or more than human.

--- Aristotle, Politics

The ability to communicate with other individuals is an essential need for any human society. The intimate relationship with others is considered the most important need for human survival, second only to food and safety (Maslow, Frager, Fadiman, McReynolds, & Cox, 1970). In every aspect of daily life, human beings live in highly social groups and interact with each other – family, friends, colleagues and even strangers on the street. Indeed, the ability to communicate effectively is so crucial to human beings that the Turing test even considered it a criterion to distinct between a human being (or an agent with human-like intelligence) and a machine (Turing, 1950). Understanding human communication therefore remains central to research in psychology, anthropology, neuroscience, sociology, political science and arts.

Simultaneously, the twenty-first century witnessed the rapid globalization and cultural integration. Digital technology (e.g., the Internet, virtual human, social robots) has promoted social communication across cultural groups (see a review of Ting-Toomey, 1994). Communicating with individuals from other cultures is fast becoming an essential but challenging task in our daily life (see Kaplan & Haenlein, 2010; Ting-Toomey, 1994, for reviews). Yet, successful cross-cultural communication is not always guaranteed (see a review of Kinloch & Metge, 2014). Case studies of international organizations including many large multinational corporations (e.g., Apple, Google, Facebook) consistently reported that employees who have different culture background do not necessarily share the same cultural values (e.g., Bond, Wan, Leung, & Giacalone, 1985; Tedeschi, Gaes, & Rivera, 1977) or communicate in the same way (e.g., Archer, 1997; Elfenbein, Beaupré, Lévesque, & Hess, 2007). Specifically, studies have shown that employees who come from Western culture (e.g., American, European) and those who come from East Asian culture (e.g., Chinese, Japanese, Korean) act very differently in training, development, rewarding and life-work relations (e.g., Paik & Teagarden, 1995; Scullion, Collings, & Gunnigle, 2007; Triandis, 1995; Tung, 1982). For example, East Asian employees are shown accepting more verbal and non-verbal insult (e.g., anger or disgust facial expressions) from their employers (e.g., Bond et al., 1985; Tedeschi et al., 1977). In contrast, Western employees are more likely to perceive these social behaviors as an inappropriate and unequal treatment (e.g., Bond et al., 1985; Tedeschi et al., 1977). Thus, cross-cultural communication between employees sometimes is challenging to achieve due to cultural barriers (e.g., Wasti, 1998), which then causes problems for these international organizations. Studies have shown that multinational corporations that pay little attention to pre-departure training on cultural conventions and cross-cultural communication skills result in higher expatriate failure rates compared to those companies who provide relevant training for their employees (Dowling, 2008; Tung, 1982). Thus, knowledge of cross-cultural communication not only contributes to the academic field but also benefits the practical applications such as to assist the pre-expatriate training in multinational corporations.

One of the most effective ways to communicate in human societies is face-to-face communication, which provides a rich set of non-verbal information (e.g., Warkentin, Sayeed, & Hightower, 1997). Besides other sources of non-verbal information, such as voice (e.g., Knapp, Hall, & Horgan, 2013; Nowicki & Duke, 1992), body gesture (e.g., De Gelder, 2006; Mahmoud & Robinson, 2011; McNeill, 2008) and touch (e.g., Bruhn, 1978; Langland & Panicucci, 1982), the human face plays a central role in face-to-face communication. The face is a rich source of non-verbal information that is used in daily interaction. It has some relatively unchangeable features such as face shape, bone structure and complexion, which can provide vital information for identity (e.g., Bruce & Young, 1986). Importantly, the face comprises one of the most powerful tools in social communication – dynamic facial expressions. Facial expressions are facial movement patterns produced by individual facial muscle groups (i.e., Action Unit, AU, Ekman & Friesen, 1978a) or more often, a combination of such facial muscle groups. The human face is equipped with a number of individual facial muscle groups (see Table 10.1 in Marieb & Hoehn, 2007 for a full list of facial muscles and their anatomic structure) and each can vary in shape and strength over time to produce an incredible diversity of facial movements. The flexibility and complexity of facial movements therefore provide a set of sophisticated social signals that can precisely deliver a rich set of social messages.

By combining individual facial movements, human face is extremely powerful in social communication. Even before six months old, infants can voluntarily produce facial expressions to share feelings with their parents, grab attention of others and influence other people (e.g., Messinger, 2002). Interestingly, infants would also use facial expressions of others as a *social reference* to appraise ambiguous circumstances. For example, Sorce, Emde, Campos, and Klinnert (1985) show that human infants seek out and use their mothers' facial expressions to help evaluate the danger in a visual cliff test, where the depth side of the cliff was adjusted to an ambiguous height so that the infants do not know whether to cross or not. When their mothers showed 'happy' or 'interested' facial expressions, most infants crossed the cliff. In contrast, when their mothers showed 'fearful' or 'angry' facial expressions, very few infants crossed the cliff. Thus, the ability to produce and recognize facial expressions is essential to humans not only for the purpose of expressing internal states but also to avoid the impending danger by seeing a fearful facial expression of others (e.g., Gray, 1987). Knowledge of facial expression is therefore vital to many core survival skills in human society such as social learning (e.g., Bandura & Walters, 1977; Goubert, Vlaeyen, Crombez, & Craig, 2011; Williams, 2002) and theory of mind (e.g., Baron-Cohen, Leslie, & Frith, 1985; Premack & Woodruff, 1978; Tager-Flusberg & Sullivan, 2000). Consequently, it is not surprising that non-verbal communication skills, including the ability to recognize and produce facial expressions, play an important role in making more friends, increasing employment opportunities and even developing romantic relationships. Studies have shown that individuals with better non-verbal communication skills are more successful in maintaining the relationship with others (e.g., Carton, Kessler, & Pape, 1999) and have higher academic and professional achievements (e.g., Nowicki & Duke, 1992; Zorn & Violanti, 1996). The tremendous importance of facial expressions in social communication becomes especially clear when learning that individuals with impaired ability to recognize and/or produce facial expressions (e.g., caused by Autism spectrum disorder, stroke or Parkinson's disease) suffer in day-to-day social life (e.g., Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Celani, Battacchi, & Arcidiacono, 1999; Hobson, Ouston, & Lee, 1988; Sprengelmeyer et al., 2003).

Due to their central role in face-to-face communication, facial expressions have fascinated psychologists and other behavioral scientists for many years. Darwin (1872/1998), Sir Charles Bell (1824) and William James (1890/2013) are considered the most influential early scientists studying facial expressions. Their work inspired the great interest of the following researchers to study facial expressions in psychology (see Goldstein, 1983 for an extensive but not exhaustive summary of the early work in psychology on facial expressions and other face-related topics) and other fields. Yet, our knowledge about facial expression originated in those theories that were proposed long before Darwin. Philosophers such as Aristotle and Descartes believe that passions in someone's soul such as 'anger,' 'fear' and 'erotic excitement' (p.808 in Aristotle, 1913/1999) are observed by their characteristic facial expressions. Thus, Aristotle and Descartes both considered facial expression as a way to read someone's soul and a link preserving the mind-body union (Aristotle, 2000; Descartes, 1989). Importantly, ancient and medieval authors were not only interested in how facial expressions are used in Western culture, but also in non-Western cultures such as India (see Shweder, Haidt, Horton, & Joseph, 2008 for a review of literature on emotional facial expressions in Western and non-Western cultures). However, Aristotle and Descartes were mainly interested in the mind-body relationship and did not aim to understand the facial expressions per se – for example, to identify which facial movement pattern is associated with a specific social message. Instead, Aristotle and Descartes only consider facial expressions as a component (of the mind and the body) in their theories. For instance, Descartes (1989) argued that there are six classes of passions in human soul – 'wonder,' 'love,' 'hate,' 'desire,' 'joy' and 'sadness.' He believed that each passion could be transmitted by the flow of spirits from the brain and resulted in changing the rest of the body including facial expressions. However, he did not attempt to demonstrate which facial movement pattern resulted from the transmitted flow of 'wonder,' 'love' and 'hate' etc.

Along with ancient philosophers, a large number of pre-twentieth-century artists were fascinated by the power of facial expressions in communicating with their audience. In his famous handbook *On Paintings* (1435/1991), Leon Battista Alberti commented that facial expressions in a painting can touch the soul of the beholder because "*I weep with the weeping, laugh with the laughing, and grieve with grieving.*" As a master in his field, Alberti was fully aware of the complexity of facial expressions and advised artists to carefully study facial expressions (p.77). Alberti was not the only artist who racked his brain to understand facial expressions. Leonardo da Vinci worked for years on the smile of his *Mona Lisa* including observing the facial expression of hundreds of people (Faigin, 2012). However, due to the obvious aesthetic purpose of these paintings, they were rarely used as stimuli in scientific research (but see Baron-Cohen, 1996, where he used portraits painted by Velazquez and Hockney as stimuli to test facial expressions of cognitive mental states including 'thinking,' 'deciding' and 'planning'). Specifically, artists tend to use obvious and theatrical facial movements, which is typically referred as 'indicating' in theatre (Faigin, 2012). In other words, facial expressions in artworks can be modified – for example, either exaggerated or understated depends on the purpose of the artist – and therefore distorted from real life. These representations might not be appropriate to be used in scientific research. In fact, Darwin tried to study facial expressions from paintings and sculptures at the beginning of his facial expression research, but later he felt disappointed by the aesthetic nature of these artworks and concluded that this source of information was not of much use to him (p.14 in Darwin, 1872/1965).

Alternatively, early researchers focused on examining the fundamental components of facial expressions – individual facial muscle groups. Researchers were astonished by the power and the complexity of facial expressions. In 1649, John Bulwer published *Pathomyotomia*, which appears to be the first substantial English-language work for demonstrating the muscular basis of facial expressions. In his work, Bulwer proposed an intuitive psychophysiological system to taxonomize facial muscles according to their locations on the face. Bulwer described a total of 41 facial muscle groups that located in head, forehead, eyebrows, eyelids, eyes, ears, nose, cheeks, lips, mouth and tongue. For each facial muscle group, he presented a detailed list of 'evoking conditions' and the 'form of movement' under these conditions. For instance, according to Bulwer's facial expression system, the condition of 'joy' evokes corrugation at corners of the eyes, cheek bulge and contractions of lips of a smile. In contrast, the condition of 'sadness' evokes wrinkled eyebrow. Importantly, it should be noted that for each facial muscle group, Bulwer described a very broad set of social messages in his list of 'evoking conditions' – not only emotions such as 'joy,' 'fear,' 'pride' and 'contempt' but also cognitive mental states such as 'meditating,' 'absent mind,' 'assent' and 'doubtful.' Bulwer considered dynamic facial expressions as "fluid signifiers of mental activity." His work published in *Pathomyotomia* was almost unique among the pre-twentieth-century face studies regarding its aim for classifying facial muscle groups according to their associated social messages (Geen & Tassinary, 2002).

Yet, Bulwer's system of facial expressions should be interpreted with caution because his theoretical foundations reflect the knowledge of physiology at his time and were much outdated in light of modern research (see Geen & Tassinary, 2002 for a review). The accuracy of his facial expression system, which was based on subjective observation of visible facial muscle movements, is therefore criticized later by other researchers (e.g., Darwin, 1872/1965; Duchenne, 1862/1990). However, as a pioneer in the field, Bulwer provided a relatively detailed blueprint to study facial expressions on their muscular basis and inspired later scientists such as Bell (1824) and Duchenne (1862/1990).

However, none of these pre-twentieth-century authors attempt to explain **why** a specific facial movement pattern is associated with a certain social message. For example, why would the wrinkled brow and the depressed mouth corners appear on the face when someone feels sad? Where does the knowledge of facial expression come from, innate or learned? Are facial expressions exclusively used for human social communication? It is already very difficult to understand facial expressions due to the complexity in their movements and associated social information. The function of facial expressions and their origins, however, are even more challenging to address. Like natural languages, the evolution of facial expressions belongs to the history so direct evidence is almost impossible (Cangelosi & Parisi, 2012). Consequently, researchers who wanted to study the origins of facial expressions must draw inferences from other indirect evidence. To understand the function and origins of facial expressions and where they come from, Darwin observed and recorded facial expressions of newborns, adults and born-blind children in a variety of cultures. He also examined the continuity of facial expressions between species by comparing human facial expressions with other animals such as gorillas, dogs and cats. His results were eventually summarized in one of his remarkable works, *the Expression of Emotions in Man and Animals*. In his work, Darwin emphasized on describing those facial movement patterns that are common across cultural groups. Consequently, Darwin's work (1872/1965) is often considered as a strong support for the universality hypothesis that is formally framed later (e.g., Ekman et al., 1987; Izard, 1994). However, it should be noted that Darwin's emphasis on cultural commonalities was due to his purpose of demonstrating the physiological origins of facial expressions. In fact, Darwin was well aware of culture diversity of facial expressions and argued that these facial expressions are likely to have evolved for social communication and influenced by cultural learning at the young age (p.16 in Darwin, 1872/1965).

To be able to understand Darwin's view of facial expressions and his contribution, it is necessary to step back and interpret his arguments from his goal of understanding the function of facial expressions. In the nineteenth century, it was widely considered that facial expressions are innate and bestowed upon humans by God. Most researchers at that time were convinced that facial movements were purely instrumental in expressions and exclusively for the purpose of human social communication (e.g., Bell, 1824; Duchenne, 1862/1990). Darwin's goal was to demonstrate the function of facial expression against this God-given knowledge assumption via two aspects. First, in line with his theory of natural selection, Darwin believed that facial expressions were the result of evolution and adaption rather than a unique gift that was exclusively bestowed upon humans. In his principle of serviceable associated habits, Darwin clearly argued that facial expressions are habitual and voluntary movements that are associated with certain mental states. Second, he disagreed with the view that the function of facial expressions is purely for social communication, which is often misinterpreted in the literature (see more discussion in Fridlund, 2014; Russell & Fernández-Dols, 1997b). In fact, his aim was to demonstrate that facial expressions are more likely originated in their physiological functions rather than for the purpose of communication. In other words, consistent with his theory of natural selection, Darwin believed that specific facial movements originated from a more direct egocentric function – to increase the chance of survival and the probability of passing on their genes to the next generation. Thus, specific facial movement patterns have evolved and persevered due to their physiological roots, which promote the lifetime reproductive success that is typically measured as ‘Darwin fitness’ or simply ‘fitness’ in later evolutionary theories (e.g., Haldane, 1927).

For example, Darwin found that humans in all cultures frown and narrow their eyes when they find something difficult or feel perplexed (p. 234). He believed that this specific facial movement originated from the time when humans were hunting outside and had difficulty seeing their prey clearly in strong sunlight. To be able to reduce the light strength and see clearly, humans have to lower their eyebrows and narrow their eyes. This physiological function of the facial muscles around the eyes – to control how much light is cast onto the retina – is further demonstrated in a recent study by Lee, Mirza, Flanagan, and Anderson (2014). Specifically, the authors showed that facial movements that produce eye widening (e.g., the facial expression of ‘fearful’ and ‘surprised’) could enhance the ability of detection. In other words, eye widening increases the amount of light gathered on the retina and therefore enhances the visual sensitivity. Conversely, facial movements that produce eye narrowing (e.g., facial expression of disgusted and perplexed) can reduce the amount of light and get light rays more focused on the retina, which enhances the ability of visual discrimination (i.e., acuity). Furthermore, Lee et al. (2014) suggested that facial expressions that produce eye widening and eye narrowing might have originated in the optical function for the differential need of enhancing detection (i.e., to see *where* is the target) and discrimination (i.e., to see *what* is the target).

It makes more sense to interpret Darwin’s theory about facial expressions in the context of biological communication (e.g., Scott - Phillips, 2008). What he suggested is that facial expressions probably started as an action with the purpose to benefit the *sender* – for example, to enhance the ability to detect a threat by widening his/her eyes. This action (i.e., the enhanced contrast of exposed eye whites) is then used as a *cue* by the *receiver* to read the sender's internal states (i.e., 'he/she is fearful.'), which can potentially benefit the receiver such as to avoid the impending threat (i.e., 'something is not safe there. I should not go there.'). At this stage, facial expression is a *cue* but not a *signal* because it provides information about internal states (e.g., 'fearful') but not for the purpose of communication. In other words, specific facial movements might have initially evolved for the purpose of benefiting the sender based on their physiological functions (e.g., to regulate sensory input) but not to influence the receiver.

However, as social activities became more common and vital for human societies through the evolution (Shariff & Tracy, 2011), some reliable facial expression cues (e.g., eye whites for fear) have then evolved as social signals through the force of habit and association (p. 369 in Darwin, 1872/1965). Different from the relatively egocentric physiological function as in the previous stage, the sender now produces facial expressions with the particular purpose to influence the receiver – for example, to change the receiver's judgments and to increase the fitness of both the receiver and the sender (Hasson, 2000). For instance, an angry facial expression can potentially benefit the receiver (e.g., 'he is angry. I should be careful') and the sender (e.g., 'I am a tough guy. Go away'). Over time, as social communication became a crucial ability for any human, the adaptive value of facial expressions might then shift from their physiological functions toward social communication (Shariff & Tracy, 2011). In other words, facial expressions could have become ritualized through human evolutionary history (Eibl-Eibesfeldt, 1979) – a process when behaviors became exaggerated and stereotyped for the purpose of social communication and more independent of its original utilitarian function (Dissanayake, 1979).

If such facial expressions have evolved as social signals, their main function is therefore to support daily communication among cultural members by simplifying interactions, averting aggression and assisting social bonding (Chapman, Kim, Susskind, & Anderson, 2009; Shariff & Tracy, 2011). Thus, cultural members might have developed specific facial expressions to support daily communication using their own cultural conventions. Such cultural conventions can diverge across cultures. For example, cultural groups have significant differences in whether individuals should express or mask their emotions in front of others – whereas Americans tend to use high arousal facial movements such as nose wrinkling to show their negative feelings towards others, Japanese tend to smile to mask their negative emotions (see more discussion of display rules in Matsumoto, 1990). Thus, cultural learning is more likely to play a role in tuning the form of such facial expressions that have evolved for the purpose of social communication (Shariff & Tracy, 2011).

In sum, if facial movement patterns have originated in their physiological function, they are more likely to converge across cultures (see also Hjortsjö, 1969; Marieb & Hoehn, 2007). In contrast, culture-specific facial expressions might have developed, learned and become cultural conventions of social communication via cultural learning – which is more likely to happen in those facial expressions that might have evolved for social communication (e.g., conversational facial expressions) as predicted by Darwin (1872/1998). Thus, understanding the cultural commonalities and specificities in modern facial expressions, including conversational facial expressions, can improve our knowledge of successful communication with individuals within and across cultures. Moreover, exploring what is common and different in facial expressions across cultures can potentially provide evidence for understanding the functions of facial expressions – a key question in the evolution of human social behaviors (e.g., Fridlund, 2014) and the nature-versus-nurture debate (e.g., Ekman, 2006). Therefore, it is not surprising that Darwin's *Expression of Emotions in Man and Animals* still remains one of the most influential books in modern science.

Despite the profound impact of Darwin's work in understanding the functions of facial expressions, it should be noted that the interpretation of his work often varies across researchers (see reviews of Ekman, 2006; Russell & Fernández-Dols, 1997a). Importantly, Darwin's perspectives of facial expressions and their cultural commonalities/specificities were probably over-simplified by the universality hypothesis (e.g., Ekman, 1972; Izard, 1992), although the latter is still considered a golden rule for many studies in the field (e.g., Ekman & Rosenberg, 1997; Keltner, Ekman, Gonzaga, & Beer, 2003; Porter & Ten Brinke, 2008; Waller, Cray Jr, & Burrows, 2008; Waters, 1987; Yacoob & Davis, 1996). For example, Ekman, Sorenson, and Friesen (1969) concluded that facial expressions of emotions are similar among humans regardless of culture (p. 87) and cited Darwin (1872/1998) as a support for their claim. However, as I stated earlier, Darwin was well aware of the broad set of social messages – not only emotions – that are associated with facial expressions. To further discuss how the universality hypothesis could possibly limit our understanding of facial expressions, I will now present a review of the existing knowledge of facial expressions in the field.

1.1 Facial Expressions of Emotions and the Universality Hypothesis

To develop a tool for measuring facial expressions and describing the facial movement patterns that can be used to distinguish different emotions, Ekman and Friesen (1978a) developed the Facial Action Coding System (FACS) and coded facial expressions of emotions accordingly. FACS is an objective coding system of facial movements that is based on anatomic structure of facial muscle groups (Hjortsjö, 1969). After training with FACS, a coder can decode complex facial expressions using individual facial muscle groups (called Action Units, AUs) and their intensity levels. For example, as shown in Figure 1-1, the two facial expressions (see *Example Image*) can be coded using the FACS as a combination of AUs (see *Action Units*) with the AU numbers and their intensity levels. Each AU number refers to an individual AU; for example, AU4D is Brow Lowerer at D intensity level (see more details in *Description* of how each AU is coded).

Example Image	Action Units	Description
	4D+9D+7C+25C	The lower face clearly shows parted lips, but it is not possible to decide for certain whether there is a *trace* of a jaw drop or not, no matter how closely one looks, and therefore, 26 is not scored. Thus, the 25 is scored with reference to a closed jaw, and the *marked* to *pronounced* difference from the 25B criteria puts its intensity at C. The *severe* raising of the medial part of the infraorbital triangle, the deepened nasolabial furrow and *severe* pulling of skin up the nose, creating a pouch above the nostril wings and *severe* to *extreme* nose wrinkling, indicate 9 at a D intensity. In the upper face, the pulling down of the eyebrows is *pronounced* to *severe*, but also, the eyebrows are pulled together *severely*, allowing AU 4 at the D intensity to be scored with 9. The eye opening is narrowed to a *pronounced* degree and together with the *pronounced* straightening and wrinkling of the lower eyelid, beyond what 9 and 4 could do, indicate a 7 at the C intensity.
	1C+2C+5C+ 25C+26C	The lower face only shows a dropped jaw, which is roughly midway between teeth together and the most drop by relaxing possible and would barely allow the tip of the tongue through, and lips separated no more than what the jaw drop might allow, indicating a score of 25C+26C. In the upper face, the lifting of the entire eyebrow and horizontal wrinkling across the entire forehead indicate the 1+2 combination, and the intensity guidelines for this combination indicate the *pronounced* raising and exposure of the upper eyelid and cover fold should be scored at the C intensity. The lifting of the upper eyelid reveals the partially covered iris in neutral now showing entirely with *slightly* more than a hairline of sclera, indicating a C intensity.

Figure 1-1. Example of coding facial expressions using Facial Action Coding System. The two facial expressions (i.e., *Example Image*) are coded as a combination of *Action Units* (i.e., individual facial muscle groups, for example, AU4 is Brow Lower; see Ekman and Friesen, 1978 for a full list of AUs) using Facial Action Coding System (as more details are described in *Description*). Adapted from Ekman and Friesen (1978).

Using this coding system, Ekman (1972) proposed prototypes of six basic facial expressions of emotions and their variants (although the meaning of 'basic' here is controversial, see Ekman, 1992; Ortony & Turner, 1990; Russell, 1994). For example, the prototypical facial expression of ‘happiness’ has Cheek Raiser (AU6) and Lip Corner

Puller (AU12). In contrast, the prototypical facial expression of 'sadness' has Inner Brow Raiser (AU1), Brow Lowerer (AU4) and Nasolabial Deepener (AU11). The authors then validated these prototypes by asking observers in different cultures to recognize these facial expressions as black-white photographs using an Alternative Force Choice (AFC) task. For example, Ekman et al. (1969) selected a small set of photographs (i.e., 20 to 30 black-white images per emotion) from over 3000 pictures that represented 'the pure display' of each emotion. In other words, only the stimuli showing the exactly same facial movement patterns as in the prototypes remained. The authors then presented this small set of photographs to observers in New Guinea, Borneo, the United States, Brazil and Japan and asked observers to select one from the six emotions (i.e., 'happy,' 'surprise,' 'fear,' 'disgust-contempt,' 'anger' and 'sadness') for each image. The authors reported that accuracy of recognition was above chance level (i.e., 16.7% in a 6-AFC task for six emotions) across cultural groups and thus concluded that facial expressions of these six emotions are universal among human societies (also see Ekman, 1972, 1984; Izard, 1969 for cross-cultural validation using the same validation procedure).

Since then, the universality hypothesis and prototypical facial expressions have been used as a standard set to test emotions (e.g., Eimer, Holmes, & McGlone, 2003; Reisenzein, Studtmann, & Horstmann, 2013; Sprengelmeyer, Rausch, Eysel, & Przuntek, 1998) and even considered as the only criteria to 'produce' emotional facial expressions in most face database. For example, Radboud faces database (Langner et al., 2010) contains images of facial expressions produced by 49 actors after practicing each facial expression with a FACS specialist for 25 minutes and each facial expression was recorded under the monitor of the FACS specialists by 'giving detailed instructions.' Not only for Western face database, the Japanese female facial expressions database (JAFFE) also asked their Japanese actors to produce exactly the same facial expression as shown in the prototypes (Lyons, Akamatsu, Kamachi, Gyoba, & Budynek, 1998).

Despite the broad impact of the universality hypothesis, it has received a number of critiques for the methods of selecting stimuli and validation procedure. These potential issues can potentially limit our knowledge of facial expressions. For example, using percentage in an N-AFC task as a criterion for detecting the signal without considering the response bias – one of the most prevalent cognitive biases in validation studies (Paulhus, 1991) – potentially increased Type I errors (see more discussion in Jack, 2013). Secondly, the chance level in a 6-AFC task was relatively low (i.e., < 20%). Thus, comparing recognition accuracy to chance level cannot sensitively reflect whether a facial expression can be accurately recognized by a given culture. In fact, by re-analyzing the data that were reported in the previous cross-cultural validation studies (e.g., Ekman et al., 1987; Izard, 1971), Russell (1994) demonstrated that there was a significant main effect of culture. Specifically, observers from Western and non-Western cultures show significantly different performance in recognizing the facial expressions of 'fear,' 'disgust' and 'anger.' Similarly, Elfenbein and Ambady (2002) applied a meta-analysis on 182 independent samples from 87 articles and demonstrated that the performance of recognizing facial expressions is significantly better within the same cultural group than across different cultural groups. As a result, the conclusion that facial expressions of emotions are universal should be interpreted with the awareness that there were cultural differences presented in the recognition performance.

More importantly, pre-selecting a relatively small set of facial movement patterns according to the prototypes that were primarily derived from Western culture might not be appropriate to characterize commonalities and cultural specificities of facial expressions. Firstly, the stimuli used in Ekman et al. (1969) were preselected from photographs that were all obtained in Western culture, which might not reflect the diversity of such facial expressions in other cultures. Secondly, only 6% of these images were selected under a scoring procedure (Ekman, Friesen, & Tomkins, 1971) that was preformed by Western observers. Such extreme degree of pre-selection – selecting 30 images from 3000 images that were produced and validated by Western observers – could potentially erase the diversity of facial movement patterns (also see discussion in Barrett & Gendron, 2016).

Alternatively, Elfenbein et al. (2007) instructed 60 Quebec and Gabon participants to produce facial expressions using 'non-standardized' procedure by showing emotion labels and dictionary definition for each emotion. The authors showed that the facial expressions of 'anger,' 'sadness' and 'surprise' have different facial movement patterns (i.e., different AUs, see Table 1 in Elfenbein et al., 2007) from prototypes and are specific to each cultural group. More importantly, the authors reported an in-group advantage on recognizing such culture-specific facial expressions. That is, a new set of participants achieved significantly higher accuracy with facial expressions from their own culture than the prototypes regardless of the ethnicity of people who produced such facial expressions in the stimuli. The authors concluded that cultural specificities of facial expressions were not represented in previously 'standardized' stimuli due to the pre-selection according to a small set of prototypes.

Although more and more recent studies showed that culture-specific facial expressions play an important role in supporting accurate communication (also see a review of Elfenbein, 2013), our knowledge of facial expressions is still mostly based on Western culture. Comparatively little is known about facial expressions in other cultures (but see Elfenbein et al., 2007; Jack, Garrod, Yu, Caldara, & Schyns, 2012; Jack, Sun, Delis, Garrod, & Schyns, 2016; Wallraven, Hur, & Shin, 2015). Given the increasing pressure of cross-cultural communication in daily life (e.g., Paik & Teagarden, 1995; Scullion et al., 2007; Tung, 1982; Wasti, 1998), understanding commonalities and cultural specificities of facial expressions and how it can impact cross-cultural communication is both essential and timely.

Additionally, the stimuli used in many previous studies were static images with extremely high intensity level of facial expressions. Facial movements might therefore have been exaggerated or distorted (e.g., Motley & Camden, 1988). Facial expressions in real life are both dynamic and continuous (e.g., Kilts, Egan, Gideon, Ely, & Hoffman, 2003; Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004; Yin, Chen, Sun, Worm, & Reale, 2008). Thus, dynamic stimuli can represent more ecologically valid facial movement patterns. Studies have shown that observers respond faster and more accurately when viewing dynamic than static facial expressions (e.g., Kilts et al., 2003; Sato et al., 2004; Yin et al., 2008). Importantly, dynamic facial expressions can capture how facial movement patterns change over time and therefore reflect the functions of such social signals – for example, early/late signal or sustained signal. For example, Jack, Garrod, and Schyns (2014) examined the early and late facial signals of six basic emotions and revealed that whereas early facial signals are more likely to deliver the core information (i.e., either approach or avoidance), later facial signals comprise the details of specific emotion (e.g., ‘surprise’ or ‘fearful,’ ‘anger’ or ‘disgust’). Consequently, using dynamic facial expressions can potentially characterize facial movement patterns more accurately than static images.

Finally and most importantly, it should be noted that Bulwer (1649) and Darwin (1872/1998) both showed that facial expressions can be used to communicate a broad set of social massages. Specifically, Darwin provided detailed description for the facial expressions of ‘thinking’ (p.232), ‘perplexed’ (p.234), ‘disagreement’ (p.290) and ‘affirmation’ (p.286) and the physical states of ‘pain’ (p.8). Ellsworth (2003) argues that facial expressions of emotions were selected and studied as prototypes not because they were the most frequently used in daily communication but were more clearly understood at that time. Since facial expressions of other social messages – for example, *conversational facial expressions* – are more likely to vary across cultures as suggested by Darwin (1872/1998), it is not surprising that emotional facial expressions were more clearly understood at that time. Certainly, it is not common for human beings to continuously experience and express strong emotions. In fact, Rozin and Cohen (2003) show that the most pervasive forms of facial expressions in real social context (e.g., a face-to-face dialogue) are those to manipulate the flow of conversations such as 'thinking' and 'confused.' Apart from conversational facial expressions, facial movements are also frequently used to communicate physical states such as physical pain and pleasure (e.g., Darwin, 1872/1998; Fernández-Dols, Carrera, & Crivelli, 2011; Goubert et al., 2011; Wall, 1978), which are vital to human survival (e.g., Wall, 2000; Walters, 1994).

However, in comparison to the numerous studies on emotional facial expressions, relatively little is know about such facial expressions. For example, what facial movements are used to communicate 'confused' and are they similar or different across cultures? Such knowledge gap then prevents our understanding of what supports successful cross-cultural communications and why it can go wrong. I will now review the current knowledge of a much smaller field of facial expressions of mental states and physical states.

1.2 Facial Expressions of Mental States and Their Function in Modulating Social Interactions

Mental states are present activity and temporary state of mind including both emotions and cognitive states such as *beliefs* and *intention* (Allport & Odbert, 1936). Since the (six basic) emotions became dominant in the field of facial expressions, researchers often use 'complex mental states' to distinguish the rest of mental states from emotions (e.g., El Kaliouby & Robinson, 2005). Studies have shown that there is a broad set of mental state terms in Western and East Asian culture. For example, Allport & Odbert used the definition of "*present activity, temporary states of mind, and mood*" and extracted 4541 terms of mental states from the second edition of Webster's Unabridged Dictionary of the English Language (1925). Wen (2007) obtained a set of 765 Chinese mental state terms that were defined as "*describing mental activity (心理活动) or mental states (心理状态) such as emotion (感情), intention (意向) and cognition (认知)*." Thus, it is not surprising that more and more studies have shown that facial expressions and/or other non-verbal communication signals including gestures (e.g., Kendon, 2004; McNeill, 2008) and vocalization (e.g., Dittmann & Llewellyn, 1968; Pammi & Schröder, 2009) are frequently used for communicating other mental states besides the (six basic) emotions.

A number of researchers have suggested the prevalence of using facial expressions of mental states to manipulate the flow of social interaction, which is often discussed as *conversational facial expressions* in the literature. For example, Birdwhistell (1970/2010) and Eibl-Eibesfeldt (1970) showed that the speakers often raise or lower their eyebrows to emphasize a point and ask the listener to 'pay attention here.' Ekman (1979) also found that the speakers use the Action Units around eyebrows (e.g., Inner Brow Raiser – AU1, Outer Brow Raiser – AU2 and Brow Lowerer – AU4) to indicate a pause (i.e., punctuation), a question (i.e., question mark) and word search (e.g., a thinking face). By linking such conversational facial expressions to speech, Bavelas and Chovil (1997) further distinguished such facial expressions could be redundant or non-redundant in terms of their functions of communication. Specifically, facial expressions are redundant when they convey the same meaning as speech. For example, speakers often display a thinking face (i.e., raised chin, tight lips, diverted gaze) at the same time when he or she says 'I am thinking.' Although such facial expressions are redundant in their meanings, they can assist listeners to understand the content of speech (Chovil, 1991). In contrast, non-redundant facial expressions convey different and sometimes contradictory information against the meaning of verbal signals. For example, the speaker could raise his lip and briefly close his eyes as if he was bored (Bavelas & Chovil, 1997) while saying 'Oh, that is interesting.' These non-redundant facial expressions provide an interesting example to study the relationship between non-verbal and verbal communication – for example, whether listeners will rely on facial expressions or words from the speaker (see more discussion of incongruent content of verbal and non-verbal communication in Argyle, Salter, Nicholson, Williams, & Burgess, 1970).

Interestingly, not only speakers but also listeners use facial expressions to control the conversation and contribute in two-way communication. However, the literature of conversation analysis focuses more on verbal signals and therefore listeners have been typically viewed as passive (see Clark, 1996 for a review). In most communication theories, listeners were treated as mute or invisible because speakers are thought to be the only one who controls the process of social interaction. This approach was largely impacted by the definition of *sender*, *receiver,* and *one-way channel* in Shannon and Weaver (1949/2015)'s theory of information transmission. The very few studies that have considered listeners only investigated the part when they (will soon) became the ones to speak – the point of turn taking (e.g., Sacks, Schegloff, & Jefferson, 1974; Schegloff, 2000). However, Clark and Wilkes-Gibbs (1986) argue that natural conversations are rarely the case of one-way communication. Instead, both speakers and listeners have to cooperate during a conversation as the listeners play an important role of *back channel* (Yngve, 1970) to achieve a successful communication. Specifically, studies have reported that listeners often produce facial expressions that could possibly change the behaviours of speakers and therefore modulate the flow of the conversation. For example, Bavelas and Chovil (1997) and Brunner (1979) analysed natural conversations and found that listeners often press their lips together to show 'suspicion' and nod with a smile to show 'agreement.' Similarly, Chovil (1991) argued that facial expressions such as brow raising, turning down mouth corners, closing eyes and pressing lips are often used by listeners as a comment to what speakers just said. Even in a speaker-dominated conversation (e.g., speakers were asked to tell a story, while listeners were not allowed to speak at all), Bavelas, Coates, and Johnson (2000) showed that if listeners do not produce any facial expressions – which was modulated by asking them to do another irrelevant task at the same time – the speakers would talk less fluently, make more pauses as if they were waiting for feedback and significantly shorten their story. Speakers would eventually feel less pleasant after the experiment. Consequently, the authors concluded that facial expressions are important for both listeners and speakers with regards to adjusting their behaviors in daily conversations.

While these previous studies provide some knowledge of using facial expressions to communicate mental states and manipulate social interaction, their methods were primarily exploratory and descriptive. Most of researchers described facial expressions based on their observation and the sample sizes were relatively small (e.g., typically less than 10 participants). Thus, it is difficult to generalize such description to the population. In addition, it remains unknown whether such facial expressions converge and diverge across cultures. Thus, one of my aims is to address this knowledge gap by objectively characterizing commonalities and cultural specificities of these conversational facial expressions. Furthermore, I would like to examine how these facial expressions are used in cross-cultural communication – for example, where they can support or hinder social interactions across cultures.

Apart from modulating the flow of social interactions, facial expressions are frequently used to communicate negative and positive states (e.g., Darwin, 1872/1998; Fridlund, 2014; Williams, 2002). Yet, there is a debate on such social messages are clearly and unambiguously signalled using facial expressions. While theories proposed that extremely negative and positive messages should be communicated by distinctive facial movement patterns (e.g., Carroll & Russell, 1996; Ekman, 1993; Russell, 1980; Schlosberg, 1952), several recent studies claim that such facial expressions of extremely are highly similar and perceptually indistinctive. I will now present a review of this debate. Specifically, I will focus on the facial expressions of two distinct physical states – the extremely negative state ‘pain’ and the extremely positive state ‘orgasm.’

1.3 Facial Expressions of Physical States and Their Discrimination on Valence

Communicating negative and positive physical states is central to human survival. For instance, facial expression of ‘pain’ can promote individual’s health and integrity by virtue of its averseness, guiding the organism away from lethal injuries and potential threats (e.g., Wall, 2000; Walters, 1994). More importantly, such intense facial expressions grab the attention of others, elicit empathic behaviour and therefore can be used a signal of asking for help (e.g., Hughes & Nicholson, 2008; Williams, 2002). In a documentary film named *A Life Without Pain*, three children with congenital insensitivity to ‘pain’ (caused by a rare genetic disorder) have to live under absolute constant parental protection since they are unaware when their hands burn from touching a hot pan or break their legs from jumping. Studies have reported that caregivers and doctors commonly use facial expressions as an indicator of whether their patients are feeling negative or positive (see reviews of Craig, Prkachin, & Grunau, 1992; Williams, 2002), or use facial expressions as diagnostic criteria to evaluate the degree of painful feeling in Western (e.g., Hicks, von Baeyer, Spafford, van Korlaar, & Goodenough, 2001) and East Asian culture (e.g., Li, Liu, & Herr, 2007). In contrast, facial expressions that communicating positive messages have been shown to contribute significantly to participant judgments of well-being (e.g., Diener, 1994; Zapf, 2002), health condition (e.g., Berry & Pennebaker, 1993; Diener & Chan, 2011) and attractiveness (e.g., Gill, Garrod, Jack, & Schyns, 2014; O’Doherty et al., 2003) in Western and East Asian culture (e.g., Dong, Jin, Cho, & Oh, 1999). One of the most important physical states – *sexual orgasm* – is associated with such fundamental aspects as sexual satisfactory (e.g., Hughes & Nicholson, 2008), which can potentially impact the sustainable development of human species via reproductive success (Meston, Levin, Sipski, Hull, & Heiman, 2004) and mate selection strategy (Morris, 1967/2010). Consequently, as one of the most efficient ways to communicate internal states in social interactions, facial expressions should convey positive and negative valences in a clear and unambiguous manner. Some researchers argue that facial expressions can be used to communicate discrete social messages (e.g., Ekman, 1992; Ekman, 1993; Etcoff & Magee, 1992; Izard, 1992; Tomkins, 1962). Others emphasize the connection between social message categories in dimensional semantic network – for example, using two dimensions such as valence and arousal (e.g., Russell, 1980), pleasantness/unpleasantness and attention/rejection (e.g., Schlosberg, 1952) or independent mono-polar factors (e.g., Clyde, 1963; McNair & Lorr, 1964; Nowlis & Nowlis, 1956). Dimensional theories argue that facial expressions can represent more than one discrete social message and the discrimination between neighboring categories (e.g., 'anger' or 'fear') relies on the contextual information. For example, Carroll and Russell (1996) show that facial expressions of 'anger' were recognized as 'fearful' when the observers were exposed to a frightening situation elicited by a scenario. Although the discrete social message theories and the dimensional theories have different emphases – either on the specific social message category or the connections between categories – both argue that the extremely positive and negative valences in facial expressions are clearly displayed and unambiguous. For example, the dimensional model of emotions (Russell, 1980) uses valence as one of the main dimensions and predicts that extremely positive and negative states should be located at two opposite ends.

However, several recent studies claim that facial expressions of extremely negative (e.g., 'pain,' 'losing') and extremely positive (e.g., 'orgasm,' 'winning') are highly similar and thus cannot support accurate communication. For example, Aviezer, Trope, and Todorov (2012) used the photographs of professional tennis players when they either won or lost high-stakes matches and asked participants to discriminate positive or negative valence from their facial expressions and/or bodies. The authors reported that facial expressions were ambiguous because participants had very similar valence ratings (both slightly negative) for winning and losing stimuli. Thus, the authors concluded that discriminating extremely positive and negative states could only be achieved using other sources of information such as the body. In addition, several studies examined photographs of pain and sexual excitement and suggested that facial expressions of pain and sexual excitement might share some facial movement patterns (e.g., Fernández-Dols et al., 2011; Hughes & Nicholson, 2008; Masters & Masters, 1986).

The conclusion that facial expressions are non-diagnostic for extremely positive and extremely negative valence is surprising because it challenges the theory that facial expressions have been evolved an effective signal in social communication (e.g., Darwin, 1872/1998; Fridlund, 2014; Smith, Cottrell, Gosselin, & Schyns, 2005). Given that distinguishing between extremely positive and negative social messages is crucial for avoiding lethal threat, the facial movement patterns should be distinct and diagnostic to optimize the transmission of such intense valence information (e.g., Smith et al., 2005).

However, there are some potential issues regarding the stimuli and the methods that were used in previous studies. Firstly, it should be noted that most studies that reported the similarity between extremely positive and negative facial expressions rely on either observation or incomplete FACS coding of the stimuli. For example, the prototypical facial expressions of 'pain' (Prkachin, 1992) used in the previous studies only include four AUs: Cheek Raiser (AU6), Lid Tighter (AU7), Upper Lip Raiser (AU10) and Eye Closed (AU43). However, a number of other AUs such as Brow Lower (AU4) and Lip Stretcher (AU27) have shown that are consistently associated with 'pain' (see Table 2 in Williams, 2002 for a detailed but not exhaustive list of AUs in the facial expressions of pain). In the similar vein, Aviezer et al. (2015) only reported four AUs – Brow Lowerer (AU4), Cheek Raiser (AU6), Mouth Stretch (AU27) and Lip Corner Puller (AU12), while there are other AUs such as Eye Closed (AU43) and Upper Lip Puller (AU10) were present in their stimuli (see Figure 1 in Aviezer et al., 2015). Thus, such argument (i.e., extremely positive or extremely negative facial expressions shared some AUs) might be misleading based on incomplete FACS coding and comparison using such a small set of AUs. It is not clear why some AUs are coded while others are not. The selection of AUs also varies across studies, which makes it difficult to assess and generalize the results. Consequently, it remains unknown whether facial expressions of extremely positive or negative states are similar or different and what facial movements are diagnostic for signalling the valence.

Secondly, it is not clear how stimuli were selected and whether or not they are representative of extremely positive and negative facial expressions. Similar to what happened in generating the stimuli for prototypical facial expressions, it is likely that the stimuli (e.g., photographs of tennis players winning or losing matches in Aviezer et al., 2012) were selected using some extent of pre-selection – for example, selecting a small set of stimuli from thousands of images on the Internet. In other words, the resulting set of images might comprise only a narrow and niche set of extremely positive and negative facial expressions that are actually used in real life. Thus, the results might be misleading if the selected stimuli were incomplete and potentially biased.

Therefore, it is essential to use an approach that could cast a wider net on the selection of such facial expressions in order to characterize their diversity more accurately. Specifically, my interest is facial expressions of negative and positive physical states – physical 'pain' and sexual 'orgasm' due to their vital roles in human evolution. Both 'pain' and 'orgasm' play important roles in any human society because they are associated with the fundamental needs of survival such as avoiding danger (e.g., Wall, 2000; Walters, 1994) and reproductive success (Meston et al., 2004). Several studies discussed the consistency in facial expressions of 'pain' across cultures (e.g., Ekman, 1989; Prkachin, 1992; Rosmus, Johnston, Chan-Yip, & Yang, 2000) but only based on a small set of AUs (e.g., four AUs in the prototypical facial expression of 'pain') and might not characterize their facial movement patterns across cultures (see a review of Williams, 2002). Even less is known about the facial expressions of 'orgasm' due to the difficulty in using methods that are permitted by ethical and cultural convention (see more discussion in Fernández-Dols et al., 2011; Hughes & Nicholson, 2008). Thus, I would aim to identify the facial movement patterns of physical pain and pleasure in Western and East Asian culture. Specifically, the first study of my thesis focused on two questions – whether physical pain and pleasure have distinct facial movement patterns (i.e., *physically* distinct) and can be discriminated by the observers (i.e., *perceptually* distinct).

Together, facial expressions are powerful and efficient in communicating a broad set of social messages and play essential roles in social interactions – showing different emotions, modulating the conversations and indicating positive or negative physical states. Yet, our knowledge of facial expressions is limited due to methods and stimuli used previously – for example, a relatively small set of facial expressions (mostly the prototypes of six basic emotions) is tested on static images and mostly based on Western observers. To improve our knowledge of facial expressions and address the new challenges of cross-cultural communication in daily life, I would aim to go beyond the current knowledge of facial expressions and investigate a broader set of facial expressions used in Western and East Asian cultures. To do that, I use a psychophysical method called reverse correlation (Ahumada & Lovell, 1971), combining it with 4D computer graphics (Generative Face Grammar, GFG, Yu, Garrod, & Schyns, 2012) and subject perception (Jack, Garrod, et al., 2012; Jack et al., 2016). Before presenting the specific studies that I conducted, I will now present an overview of why and how I use these methods to investigate facial expressions.

1.4 The Data-Driven Approach: Sampling Facial Movement Patterns

The universality hypothesis and prototypical facial expressions (e.g., Ekman & Friesen, 1978a) have been considered a gold standard for a long time. However, as I presented earlier, traditional methods of generating stimuli by selecting a small set of images according to the prototypes that were primarily based on Western cultures might not necessarily characterize the diversity of facial expressions in other cultures (see Elfenbein, 2013; Jack, 2013; Russell, 1994, for reviews). To address this problem, several studies recorded and analyzed spontaneous facial expressions – facial expression that is not produced by direct, external requests by another person (e.g., Ekman & Rosenberg, 1997; Matsumoto & Willingham, 2006; Wagner, MacDonald, & Manstead, 1986). To obtain spontaneous facial expressions, some studies used the photographs captured in the real life scenarios as the stimuli. For instance, Matsumoto, Olide, Schug, Willingham, and Callan (2009) used facial expressions of Olympic judo athletes and reported that American and Japanese participants show agreement when categorizing the emotions of these spontaneous facial expressions using an AFC task with alternatives of six basic emotions, neutral and other. Other studies recorded facial expressions of the participants when they viewed emotionally loaded images (e.g., Wagner et al., 1986) or films (e.g., Philippot, 1993). However, recording and analyzing spontaneous facial expressions is also currently problematic because facial expression elicitation methods – either in real life scenarios or by emotionally loaded stimuli – can generate combinations of internal states. For example, Gross and Levenson (1995) argue that an emotionally loaded film usually elicits a few emotional states simultaneously. A disgusting film can elicit 'horror,' 'anger,' and 'surprise' and correspondingly complex facial expressions. As a result, accurately measuring the internal states of participants at each time point is virtually impossible. Consequently, it is often suggested that human observers – experts of categorizing facial expressions – could be used to interpret or validate the produced facial expressions (e.g., Russell & Fernández-Dols, 1997a; Smith et al., 2005; Susskind, Littlewort, Bartlett, Movellan, & Anderson, 2007).

Importantly, human observers rely on diagnostic features – features that are relevant to the task (e.g., to determine whether the facial expression displays 'happiness') – to represent and categorize social message categories (e.g., Gibson, 1969; Lawrence, 1949; Schyns, Goldstone, & Thibaut, 1998). The knowledge of such features is stored in participant's prior knowledge that was acquired biologically and/or through experience and learning (see Chisholm, 1966; Murphy, 2004, for more discussion on the structure of knowledge and concepts). In other words, when observers detect a given social message – for example, whether a facial expression is happy or not, they look for specific feature that matches their mental representation of that category – that is, a perceptual category reflecting their understanding of what a happy facial expression look like based on their prior knowledge. For instance, Smith et al. (2005) used the Bubbles method (see more details in 1.4.1) to show that smile is the diagnostic feature for facial expressions of 'happiness,' whereas wide open eyes are diagnostic for the facial expression of 'fear.' Therefore, if a face stimulus has the feature that matches participant's knowledge of happy facial expressions (e.g., a smile), this face stimulus will be accurately categorized as 'happy.' In contrast, if a face stimulus does not have the feature that matches participant's knowledge – either no smile or a 'different' smile (e.g., a bitter smile, Marcus, 1982), observers will have significantly lower accuracy when categorizing such facial expressions. Therefore, understanding how facial expressions are used to communicate social messages requires the knowledge of what facial movement patterns (i.e., individual AU or the combination of AUs) are used for communicating a specific social message. The mismatch between the information presented in stimuli and participant's knowledge of facial expressions can potentially happen if the stimuli are strictly selected using a theory-driven approach, for example, selecting a small set of images that only display a given facial expression. If a facial movement is not predicted by the theory, such facial expressions will therefore not be present in the stimuli. For example, Elfenbein et al. (2007) reported that Quebecois participants consistently associated Brow Lowerer (AU4) and Lid Tightener (AU7) with '*contempt*,' which were not presented in the prototypical facial expressions of '*contempt.'* Unfortunately, due to the dominance of the universality hypothesis, little is known about such culture-specific facial expressions (but see Elfenbein, 2013; Jack, Garrod, et al., 2012) and therefore existing theories of facial expressions are primarily based on Western culture. Studies have shown significantly lower accuracy for recognizing prototypical facial expressions in non-Western cultures (e.g., see Elfenbein & Ambady, 2002; Nelson & Russell, 2013, for reviews), suggesting that the features used for communicating social messages in these non-Western cultures might not present in such prototypical facial expressions. Thus, using theory-driven approaches may potentially limit our knowledge of commonalities and cultural specificities of facial expressions.

Instead, using data-driven approaches may characterize diversity of facial expressions more accurately. Data-driven methods such as reverse correlation (Ahumada & Lovell, 1971), Singal Detection Theory (Green & Swets, 1966) and ideal participant analysis (Geisler, 1989) are typically used in the field of psychophysics to uncover the 'law' between objectively measurable stimuli (e.g., the wavelength of light) and subjective perception of observers (e.g., color categories such as red). To investigate the relationship between physical stimuli and participant's behavior, psychophysicists often quantitatively (in most case, parametrically) manipulate the properties of physical stimuli (e.g., wavelength of light, contrast, spatial frequency) and measure how participant's behavior (e.g., perception of color categorization, color brightness) changes accordingly. The relationship between the changes of stimuli and participant's behavior is then systematically examined, for example, using the reverse correlation method (Ahumada & Lovell, 1971). Thus, a psychophysics law can precisely identify what information is used for the participant's perception – for example, which range of light wavelength is perceived as red. Using such psychophysical approaches I can then precisely understand what facial movements are used for communicating a specific social message category – a question that is often not clearly addressed by the traditional approaches in the field of facial expressions. In addition, such psychophysical approaches are typically data-driven. In other words, within biological/physical limits or theoretical boundaries, the measurable dimension of stimuli is sampled agnostically and the whole range of a given behavior is systematically tested in the experiment. For example, within the range of luminance that can be perceived for human vision – between 10^{-6} and 10^{6} cd/m^2 (see Schubert, 2006), the luminance of stimuli was parametrically manipulated and the whole perception of brightness spectrum was tested from no perception to the maximum perception threshold (e.g., Livingstone & Hubel, 1987). Compared to the traditional theory-driven approaches – which test a relatively small set of prototypical facial expressions – such data-driven methods can test a much broader set of facial movement patterns and identify the facial movement patterns that support the accurate social communication. Therefore, such data-driven methods can enhance our knowledge by characterizing diagnostic feature and diversity of facial expressions.

However, as I presented earlier, facial expressions are combinations of a number of individual facial movements (i.e., AUs) and each can vary in its intensity level and form of movements. Thus, systematically sampling human facial movement patterns is much more complicated than changing the luminance level. I will now present a review of three main methods that have been used to sample facial movement patterns in the literature – systematically sampling the features from facial expression images (i.e., 1.4.1), adding pixel noise to a neutral face (i.e., 1.4.2) and sampling individual Action Units (1.4.3).

1.4.1 The 'Bubbles' Method: Sampling Features From Facial Expressions

To systematically sample information used in recognition tasks and reveal what information is task-relevant, Gosselin and Schyns (2001) developed the Bubbles method, which adds mask on the base images. For example, images of neutral and happy facial expressions, punctured by a number of randomly located Bubbles (i.e., smooth and symmetrical Gaussian windows). Thus, information of the face is sampled by randomly located Bubbles. Figure 1-2a shows an example. Each stimulus therefore contains the information, just by chance, relevant or not relevant to the task. On each experiment trial, observers viewed a stimulus and responded accordingly – for example, to determine whether the presented facial expression was happy or not. If the stimulus represents task-relevant information (e.g., a smile), observers can recognize that the stimulus displays a happy facial expression. In contrast, if the stimulus does not represent task-reverent information, observers are less likely to recognize happy facial expression from the stimulus. After many such trials, the stimuli that elicited correct responses (i.e., here, when observers recognized the happy facial expression) can be added together to reveal what information is consistently used in the task (i.e., here, which facial movement patterns are used to recognize the happy facial expression) by correlation, linear regression (e.g., Schyns, Bonnar, & Gosselin, 2002; Smith et al., 2005) or Mutual Information (e.g., Rijsbergen, Ince, Rousselet, Gross, & Schyns, 2016). Therefore, the Bubbles method is a powerful approach to precisely understand what facial movement patterns are used in communicating a specific social message category. However, since the Bubbles were added to a base image, it can only sample information from the base image. In other words, the space of sampling is limited by the base image. If a facial movement pattern is not present in the base image, the Bubbles method therefore cannot reveal whether this facial movement pattern is task-relevant or not. As I presented earlier, the stimuli set used previously (e.g., base images of emotional facial expression) might not characterize the broad set of facial expressions in daily communication. As a result, sampling information from these stimuli may still not improve our knowledge beyond emotional facial expressions. To address this problem, an alternative approach is to sample information without the limit of the base images– for example, by adding pixel noise to a face image with neutral facial expression. I will now present a brief introduction of this approach.

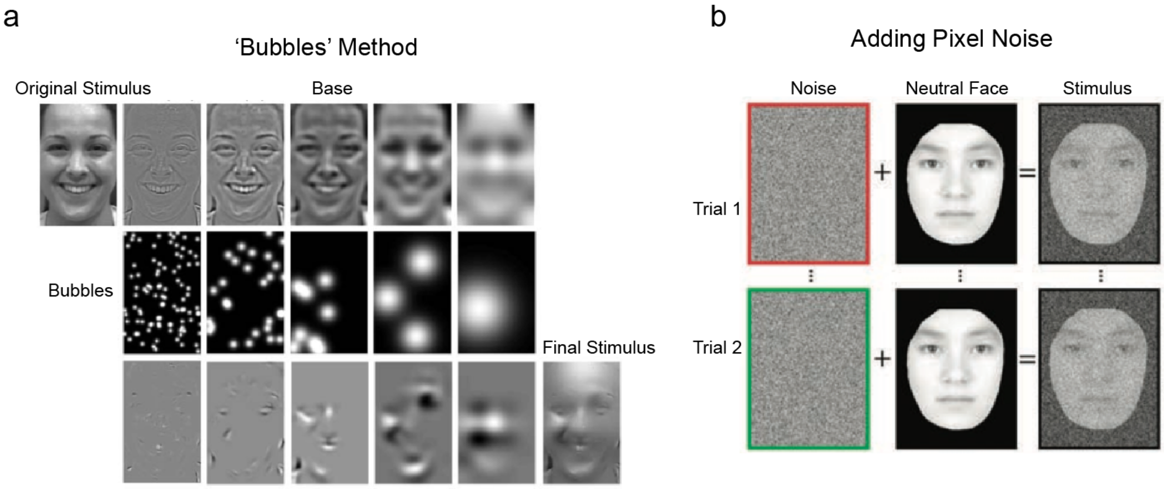

Figure 1-2. Illustration of reverse correlation methods. (a) *'Bubbles' method:* this shows an example of adding bubbles to the top left image (i.e., 'Original Stimulus') in Smith et al. (2005). Specifically, each original face image was decomposed to five spatial frequency bandwidths (i.e., 'Base') and then each added bubbles (i.e., randomly positioned Gaussian windows). Finally, five base images were combined and generated the final stimulus. (b) Adding pixel noise: this shows an example of adding white noise (i.e., noise) to a neutral face image in Jack, Caldara, and Schyns (2012). Specifically, white nose was added to generate stimuli that can elicit perception of emotional facial expressions – for example, trial 1 (color-coded in red) may elicit the perception of 'sad,' whereas trial 2 (color-coded in green) may elicit the perception of 'anger.' Adapted from Smith et al. (2005) and Jack, Caldara, et al. (2012).

1.4.2 Adding Pixel Noise To Face Stimuli

To sample information beyond existing facial movement patterns of a base image, Jack, Caldara, et al. (2012) added uniform white noise to face images displaying neutral facial expression. Figure 1-2b shows an example. The resulting stimuli can potentially change the observers' perception of the facial stimuli because the face stimuli with added white noise can form a random pattern that, just by chance, comprises the task-relevant information. For example, if the white noise is superimposed on the eye whites of the face stimulus in one experiment trial, observers may potentially categorize this face stimulus as 'fearful' because the added white noise elicited the perception of wide opening eyes, which is diagnostic for a 'fearful' facial expression. Similar to the Bubbles method, the stimuli that elicited a given behaviour (e.g., observers categorized the stimuli as 'fearful') can be added together to compute a *Classification Image* (Ahumada, 2002) – an image that shows the task-relevant information (see Figure 1-2b). Since the information pattern resulting from adding the white noise is relatively independent of the base image (i.e., the neutral face), using such a method can potentially reveal the task-relevant information even if it is not presented in the based image. For example, Jack, Caldara, et al. (2012) showed that East Asian observers require more information around the eyes to perceive emotions, whereas Western observers primarily rely on the mouth.

However, since each added pixel is considered an independent parameter in the sampling procedure, such method requires a very large number of trials to reveal a statistically robust classification image. For example, Jack, Caldara, et al. (2012) used 12,000 trials for each participant. Van Rijsbergen, Jaworska, Rousselet, and Schyns (2014) used a similar method to investigate the perception of the age of the face, where each participant completed 3,240 trials. In addition, as I presented earlier, static images of facial expressions may not represent the richness of dynamic facial expression in real life (e.g., Back, Jordan, & Thomas, 2009), which may still limit the information space that can be sampled using such data-driven approach.

A better way is therefore to directly sample the fundamental components of dynamic facial expressions – that is, the individual face movements (i.e., AUs) and their combinations. Achieving this sampling requires precise controlling of individual AU movement and flexibly generating the combination of these AUs. To do this, I used a 4D computer graphic platform (Generative Face Grammar, GFG, Yu et al., 2012), which comprises a library of over 40 AUs and can precisely control the movement of each AU.

1.4.3 The Generative Face Grammar

The GFG (Yu et al., 2012) is a platform that can flexibly synthesize dynamic facial expressions by parametrically modulating the movement of the fundamental components of facial expression – the individual face movements (i.e., AUs) and their combinations. The GFG comprises a library of over 40 core Action Units and controls the movement of each AU using 7 parameters (i.e., onset, acceleration, peak latency, peak amplitude, deceleration, offset, sustain). Specifically, the GFG was developed by recording each of core AUs that was produced by four different FACS coders using stereo-optic equipment. Thus, for each individual AU, this procedure produced a 3D shape and texture model. By mapping the model between each AU template and a neutral face captured using the same procedure, the GFG can then parametrically control the movement of each Action Unit on any face (see Yu et al., 2012, for more details). Therefore, the resulting dynamic and photo-realistic 3D face stimuli generated by the GFG can provide a rich information space, which is an ideal base for sampling the dynamic facial expressions used in data-driven methods. In addition, since the resulted dynamic facial movement patterns can be mapped on any face, the GFG provides a possibility of easily transferring these dynamic facial expression models to the digital agents (e.g., social robots and virtual humans), which will therefore improve their interactive skills to engage with their human users (e.g., Cassell, Sullivan, Prevost, & Churchill, 2000).

In sum, facial expressions play a central role in human social communication by communicating a broad set of social messages – including emotions, mental states and physical states. Understanding which facial movement patterns are used for communicating a specific social message category is essential for understanding the physiological and social functions of facial expressions. Yet, little is known about facial expression beyond the six basic emotions and the Western culture. Using a data-driven approach can provide a better understanding of a broad set of facial expressions and demonstrate how a given specific facial movement pattern drives our social judgments. In addition, this approach can characterize the commonalities/diversities of such facial movement patterns across cultures.

To explore how facial movement patterns are used to communicate distinctive social messages, I first started my research on the facial expressions of the diametrically opposite messages – physical pain and pleasure (Study 1). To further understand the relationship between facial movement patterns and their corresponding social messages, I then expanded my research to a broad set of mental states in Western and East Asian culture (Study 2). Finally, I examined the role of culture in impacting cross-cultural communication clarity using the four key conversational facial expressions (Study 3). Together, my research examined how specific facial movement patterns are used to communicate a broad set of social messages in Western and East Asian culture and how culture shapes the signalling and perception of facial expressions.

2 Study 1: Examining Physical and Perceptual Distinctiveness of Facial Expressions of Pain and Pleasure

2.1 Introduction

Humans are one of the most culturally diverse species, yet one constant across all humans is the experience of physical pain and pleasure (e.g., Frijda, 2007), which are associated with the most fundamental needs of human survival. For example, communicating states of pain can alert others to a potential threat (e.g., Williams, 2002) or elicit empathic behaviour (e.g., Anand, Rovnaghi, Walden, & Churchill, 1999; Prkachin, Solomon, Hwang, & Mercer, 2001). In contrast, identifying physical pleasure such sexual arousal may increase the sexual motivation of the participant and therefore promote the likelihood of future copulations (e.g., Masters & Masters, 1986; e.g., Meston et al., 2004; Symons, 1980). Therefore, accurately discriminating between pain and pleasure is advantageous to human survival (Fridlund, 2014; e.g., Hughes & Nicholson, 2008).

The experience of extreme physical pain or pleasure is often accompanied by a facial expression. One theory about the function of the facial expression is that they serve a regulatory function that benefits the expresser. For example, the gate-control hypothesis suggests that non-painful sensation asserts controlling painful feeling and therefore inhibits overwhelming experiences (e.g., Melzack & Wall, 1967; Wall, 1978). If facial expressions of pain and pleasure both serve the physical function of regulating overwhelming experiences, I would predict that they would be represented using similar facial signals. Indeed, many studies claim that facial expressions of extreme negative and positive states are highly similar (e.g., Fernández-Dols et al., 2011; Hughes & Nicholson, 2008; Masters & Masters, 1986) and suggests that these extreme negative and positive states can only be communicated using other sources such as the body (e.g., Aviezer et al., 2012). Yet, other researchers argue that these facial signals have evolved for social communication. For example, sexual coitus is fundamental to survival of the species. One of the most salient, unique, and universal features of human coitus is face-to-face interaction, suggesting that facial expressions of sexual pleasure might play a central communicative role (e.g., Baumeister & Bratslavsky, 1999; Eibl-Eibesfeldt, 1989). With respect to painful experiences, Williams (2002) argues that the evolutionary function of the expression of pain is to demand attention and to prompt others' help to escape, recover, and heal. Biological signalling accounts predict that to clearly communicate the diametrically opposite concepts of pain and pleasure, they should be represented by distinct patterns to optimise discrimination (e.g., Darwin, 1872/1998; Smith et al., 2005). Therefore, if these facial expressions have evolved for communication, their facial signals should be distinctive and may vary across cultures (Shariff & Tracy, 2011).

To explore this, I mathematically modelled the dynamic facial movement patterns of extreme negative (i.e., pain) and extreme positive (i.e., orgasm) facial expressions in each of 40 individual observers in Western and East Asian culture (see *Experiment 1, Participants*). Following a within-cultural validation of the resulting facial expression models (see *Experiment 2*), I analysed the facial expression patterns using information theoretic approach to show that facial expressions of pain and pleasure are highly distinct. A cross-cultural analysis showed that while pain facial expressions are highly similar across cultures, facial expressions of pleasure form distinct culture-specific clusters based on culture-specific accents.

2.2 Experiment 1: Modelling Facial Movement Patterns of Pain and Pleasure.

To characterize facial movement patterns used to communicate pain and pleasure across cultures, I used a data-driven approach that agnostically samples face movement patterns and tests them against a participant's mental representations.

2.2.1 Method

2.2.1.1 Participants

I recruited 80 observers as follows – 20 Western white Caucasian females (20 European, mean age 22 years, SD = 2.67 years), 20 Western white Caucasian males (20 European, mean age 22 years, SD = 2.75 years), 20 East Asian females (20 Chinese, mean age 23 years, SD = 1.81 years) and 20 East Asian males (20 Chinese, mean age 24 years, SD = 1.81 years). To control for the possibility that the participant's mental representations could have been influenced by cross-cultural interactions, I recruited observers with minimal exposure to and engagement with other cultures (De Leersnyder, Mesquita, & Kim, 2011) as assessed by questionnaire (see *Appendix A, Screening Questionnaire*). I also recruited observers who are sexually active (as per self-report) and identified as heterosexual as assessed by the Kinsey scale (Kinsey, Pomeroy, & Martin, 1948). All East Asian observers had arrived in the UK for the first time with an average UK residence of 3 months (SD = 1.3 month), and possessed a minimum International English Testing System (IELTS) score of 6.0 (Competent user). All observers had normal or corrected-to-normal vision, and were free from any emotion related atypicalities (e.g., Autism Spectrum Disorder, depression, anxiety), learning difficulties (e.g., dyslexia), synaesthesia, and disorders of face perception (e.g., prosopagnosia) as per self-report. I paid each participant £6 per hour, and obtained their written informed consent. The University of Glasgow College of Science and Engineering Ethics Committee authorized the experimental protocol (Ref ID: 300140074).

2.2.1.2 Stimuli

Figure 2-1 illustrates the stimuli and procedure using one illustrative trial. On each experimental trial, a dynamic facial expression generator – the Generative Face Grammar (GFG; Yu et al., 2012) – randomly selects a biologically legitimate combination of individual facial movements called Action Units (AUs; Ekman & Friesen, 1978b) from a core set of 42 AUs (minimum = 1, maximum = 6, median = 3). For example, on this illustrative trial in Figure 1 three AUs are randomly selected – Brow Lower (AU4) color-coded in blue, Nose Wrinkler (AU9) color-coded in green and Lip Stretcher (AU20) color-coded in red. A random movement is then assigned to each individual AU using randomly selected values for each of seven temporal parameters (i.e., onset latency, acceleration, peak amplitude, peak latency, sustain, deceleration, offset latency; see labels illustrating the blue curve). The randomly activated AUs are then combined to produce a photo-realistic random facial animation (duration 1.25 seconds) shown here with four snapshots across time. For each of the four participant groups (20 males, 20 females in each culture), I displayed each of 3600 randomly generated facial animations on one of 4 same-race, sex-opposite face identities (white Caucasian females: mean age 26 years, SD 3.95 years; white Caucasian males: mean age 29 years, SD 3.59 years; East Asian females: mean age 26 years, SD 3.27 years; East Asian males: mean age 27 years, SD 3.46 years). I generated all stimuli using a standard procedure in 3D Studio Max (see Yu et al., 2012 for full details).

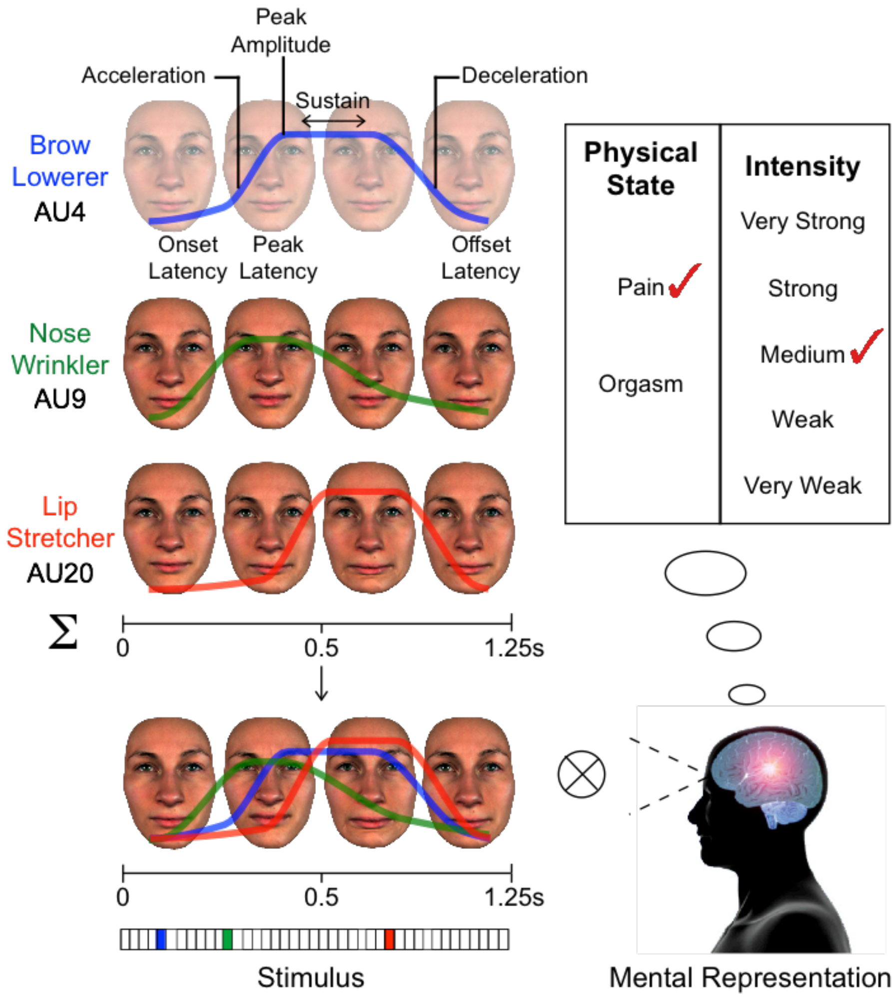

Figure 2-1. Modelling dynamic mental representations of facial expressions of physical pain and pleasure. *Stimulus.* On each experimental trial, a dynamic facial expression generator (Yu et al., 2012) randomly selects a biologically legitimate combination of individual facial movements called Action Units (AUs; Ekman & Friesen, 1978b) from a core set of 42 AUs (here, Brow Lower – AU4 color-coded in blue, Nose Wrinkler – AU9 in green, and Lip Stretcher – AU20 in red). A random movement is then assigned to each AU separately using randomly selected values for each of seven temporal parameters (i.e., onset latency, acceleration, peak amplitude, peak latency, sustain, deceleration, offset latency; see labels illustrating the blue curve). The randomly activated AUs are then combined to produce a photo-realistic random facial animation, shown here with four snapshots across time. The color-coded vector below shows the three AUs randomly selected on this example trial. *Mental Representation.* The participant views the facial animation and if the random face movement pattern correlates with their mental representation of either pain or pleasure, they categorize it accordingly (here, 'pain') and rate its perceived intensity on a 5-point scale from 'very weak' to 'very strong' (here, 'medium'). Otherwise, the participant selects 'other.' Each of 40 observers in two cultures – Western and East Asian – categorized the same 3600 facial animations displayed on same-race, sex-opposite faces presented in random order across the experiment.

2.2.1.3 Procedure

Participants viewed the facial animation and, if the facial expression pattern correlated with their mental representation of 'pain' or 'orgasm' they categorized it accordingly (e.g., in Figure 1, 'pain') and rated the perceived intensity on a 5-point scale from 'very weak' to 'very strong' (e.g., in Figure 1, 'medium'). Otherwise, the participant selected 'other.' Each participant categorized the same 3600 animations as illustrated in **Figure 2-1**, with all stimuli presented in random order across the experiment for each participant. I presented all stimuli on a black background displayed on a 19-inch flat panel Dell monitor with a refresh rate of 60 Hz and resolution of 1024 × 1280. Stimuli (mean size 19.12 cm [SD 0.90 cm] × 13.07 cm [SD 0.86 cm]) appeared in the participant's central visual field and played only once for a duration of 1.25 seconds. A chin rest ensured a constant viewing distance of 77 cm, with images subtending 14.16° (vertical) × 9.70° (horizontal) of visual angle, reflecting the average size of a human face (Ibrahimagić-Šeper, Čelebić, Petričević, & Selimović, 2006) during natural social interaction (Hall, 1966). I presented all labels (mean size: 0.96 cm [SD 0.05 cm]) × 2.63 cm [SD 0.44 cm]) in lower case white font ('MS Sans Serif') and subtended 0.72° × (vertical) × 1.96° (horizontal) of visual angle.

2.2.1.4 Results

Following the experiment, I modeled for each individual observer their dynamic mental representations of facial expressions of pain and pleasure using an established model fitting procedure as follows. First, I performed a Pearson correlation between two binary vectors: the first detailed the presence or absence of each AU on each trial and the second detailed the response of the observer (e.g., 'pain' = 0, 'orgasm' = 1). For all significant correlations (two tailed $p < .05$), I assigned a value of 1 (0 otherwise), thus producing a 1 × 42-dimensional binary vector detailing the composition of AUs that are significantly associated with each facial expression type for that observer. To model the dynamic components of each AU, I performed a linear regression between the binary response variables and the 7 temporal parameters of each significantly correlated AU as detailed on each trial. To calculate the intensity gradients of the facial expression patterns, I fitted a linear regression model to the temporal parameters of each AU and the observer's intensity ratings. I then made the resulting dynamic facial expression patterns into movies for later use as stimuli by combining the significantly correlated AUs with their temporal activation parameters, using only the 'high intensity' ratings as these comprise the most salient signals. Thus, Experiment 1 produced a total of 160 dynamic facial expression patterns (40 observers × 2 cultures × 2 pain/pleasure) each represented as a 1 × 42-dimensional binary vector detailing the significant AUs, plus 7 additional values detailing the temporal dynamics of each significant AU. Our approach therefore delivers the precise dynamic facial expression patterns associated with pain and pleasure in each culture, which can be used as stimuli to test their perceptual distinctiveness and which can be objectively analyzed to test their physical distinctiveness. Figure 2-2 shows the distribution of AUs using (a) color-coded matrixes and (b) face maps, where the most frequent AUs are coded in red and the least ones are coded in blue (see colorbar to right).

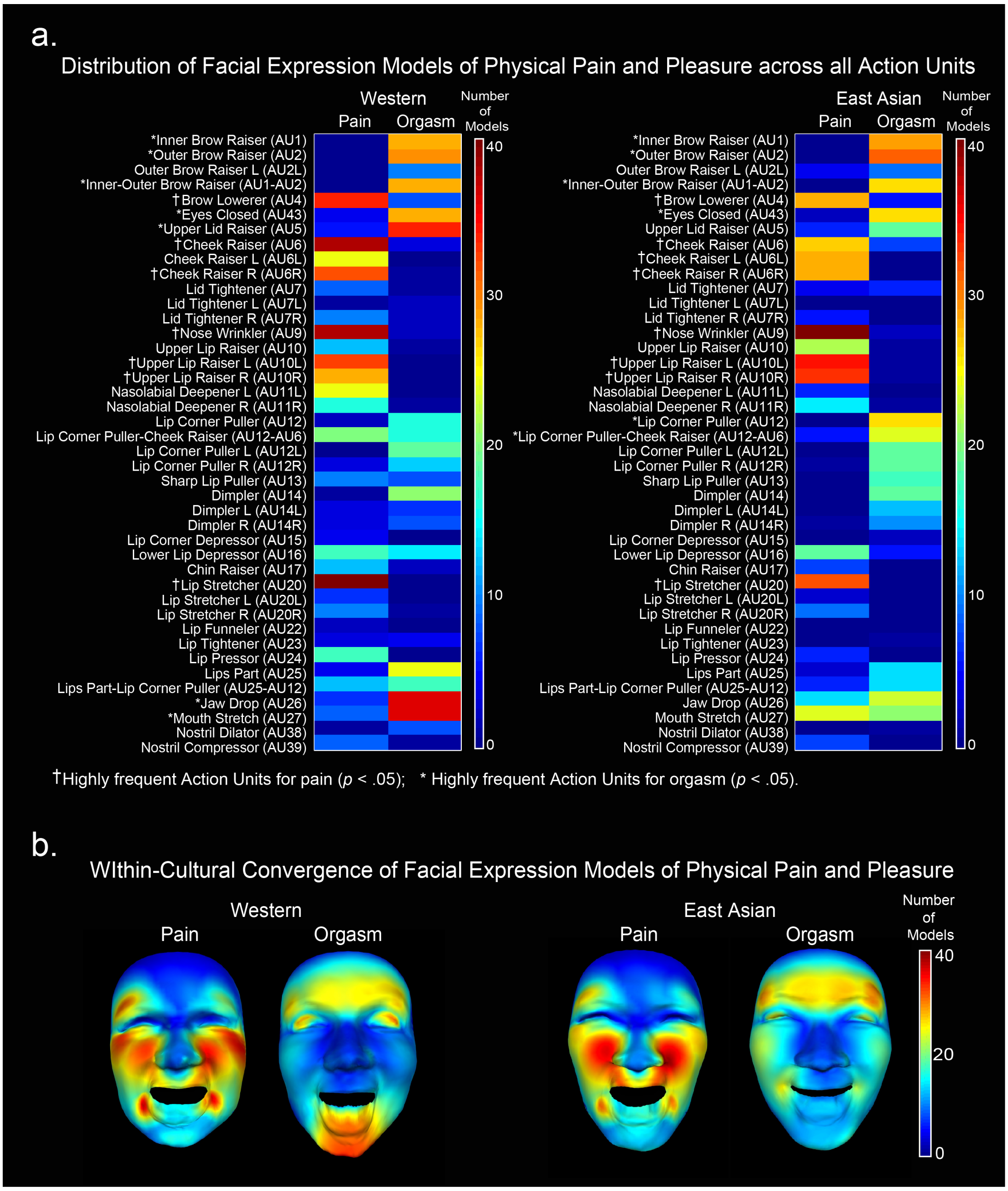

Figure 2-2. Distribution of facial expression patterns of pain and pleasure across all Action Units and their convergence across observers within each culture. (a) *Distribution of facial expression models of pain and pleasure across all Action Units.* For each culture – Western (left panel) and East Asian (right panel) – the color-coded matrix show for pain and pleasure (see labels above) the number of individual observer facial expression patterns (maximum = 40) with each Action Unit (AU; see labels on left). Warmer colours indicate more patterns; cooler colours indicate fewer patterns (see colorbar on right). For example, all Western 'pain' all facial expression patterns include Lip Stretcher (AU20) as indicated by the red colouring. In contrast, none of the Western 'pleasure' facial expression patterns include Lip Stretcher (AU20) as indicated by the blue colouring. (b) *Within-cultural convergence of facial expression patterns of pain and pleasure.* Color coded face maps show for each culture and pain and pleasure separately, the number of individual observer facial expression patterns with each AU using the same color-coding as in panel (a).

2.3 Experiment 2: Measuring Perceptual and Physical Distinctiveness of Dynamic Facial Expressions of Pain and Pleasure.

Biological signalling accounts predict that the diametrically opposite concept of pain and pleasure should be represented using highly distinct patterns. Here, I test this prediction by measuring both the perceptual and physical distinction of dynamic facial expressions of pain and pleasure in each culture separately.

2.3.1 Perceptual Distinctiveness Between Facial Expressions of Pain and Pleasure.

First, to measure the perceptual distinctiveness of pain and pleasure, I used a signal detection approach that provides a reliable measure of perceptual discriminability called d-prime (e.g., Green & Swets, 1966; Stanislaw & Todorov, 1999). D-prime reflects the signal sensitivity as it measures the distance between hit rates (signal present, 'yes' response) and false alarm rates (signal absent, 'yes' response).

2.3.1.1 Participants

I recruited a new set of 104 observers as follows – 26 Western white Caucasian females (26 European, mean age 22 years, SD = 2.22 years), 26 Western white Caucasian males (26 European, mean age 23 years, SD = 3.31 years), 26 East Asian females (26 Chinese, mean age 23 years, SD = 1.49 years) and 26 East Asian males (26 Chinese, mean age 23 years, SD = 1.67 years). I recruited observers using the same criteria as detailed above. All East Asian observers had an average UK residence of 3 months (SD = 2.4 months).

2.3.1.2 Stimuli

On each experimental trial, observers first viewed a word ('pain' or 'orgasm') displayed on-screen for 1.5 seconds followed directly by either a correctly or incorrectly matched facial expression displayed once for 1.25 seconds. I asked observers to indicate whether or not the word accurately described the facial expression by pressing 'yes' or 'no' keys on a keyboard and to respond as accurately as possible. I assigned 'yes' and 'no' keys to separate hands and counterbalanced key assignments across observers. Half of the trials comprised correct word-facial expression matches and included all 400 facial animations, with the other half of the trials comprising incorrect word-facial expression matches. Each observer therefore completed 800 trials (400 facial expressions × correct/incorrect matches). I presented all stimuli in the observer's central visual field with all stimuli presented in random across the experiment for each observer. I presented all stimuli (mean size of facial expressions = 19.21cm [SD 0.74 cm] × 13.02 cm [SD 0.73 cm]) on a black background using a 19-inch flat panel Dell monitor with a refresh rate of 60 Hz and resolution of 1024 × 1280. I presented words (mean size = 5.5 cm [SD 0.14 cm] × 2 cm [SD 2.12 cm]) in lower case white font ('MS Sans Serif') and in English. A chin rest ensured a constant viewing distance of 77 cm, with faces subtending 14.22° (vertical) × 9.67° (horizontal) of visual angle, which reflects the average size of a human face (Ibrahimagić-Šeper et al., 2006) during typical social interaction. Text subtended 4.26° (vertical) × 1.55° (horizontal) of visual angle.

2.3.1.3 Results

Following the experiment, I examined the perceptual discriminability of each individual observer facial expression pattern in each culture by computing its d-prime value using the pooled responses of all observers. Table 2-1 shows the average d-prime values across facial expression patterns for pain and pleasure in each culture.

Table 2-1

Perceptual Distintiveness (d-prime) Between Facial Expressions of Pain and Pleaure

Culture	Western	East Asian
Facial Expression	*Mean (SE)*	*Mean (SE)*
Pain	1.37 (0.08)	1.37 (0.08)
Pleasure	1.34 (0.05)	1.32 (0.11)

Figure 2-3 shows individual facial movement patterns ranked by their d-prime values in each culture.

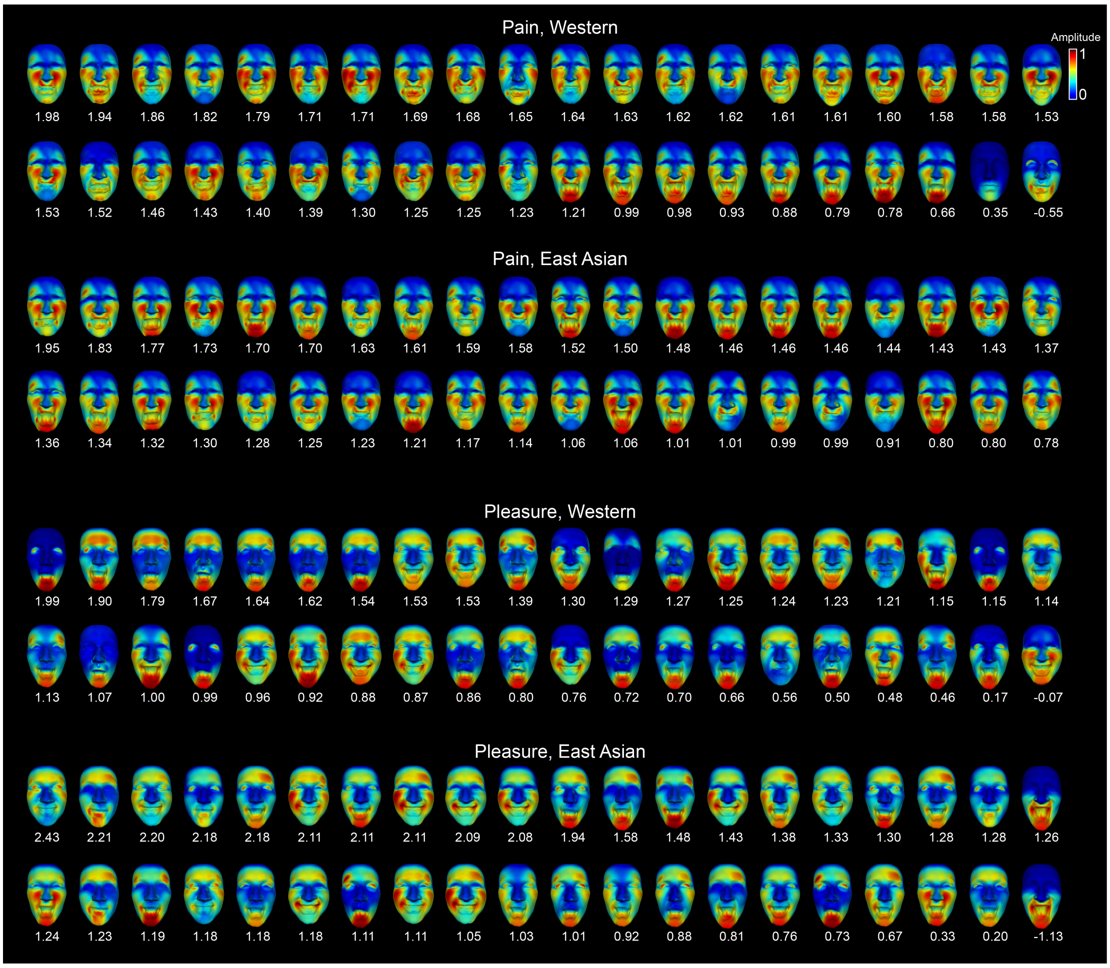

Figure 2-3. Individual facial expression patterns of pain and pleasure ranked by d-prime in each culture. Color-coded face maps show in each culture (see labels) the individual observer facial expression patterns of pain and pleasure. Color-coding represents the amplitude (i.e., intensity) of each significant Action Unit (see colorbar on right) as derived from the reverse correlation analysis. Facial expression patterns are ranked according to d-prime (below each face map) in descending order from left to right.

Previous studies suggest that the facial expressions of pain and pleasure are ambiguous to observers (e.g., Aviezer et al., 2015; Aviezer et al., 2012; Barrett, Mesquita, & Gendron, 2011), which predicts that for the hit rate equals the false alarm rate for each facial expression model. To visualise the perceptual distinctiveness between facial expressions of pain and pleasure, I plotted hit rate (signal present, ‘yes’ response) and false alarm rate (signal absent, ‘yes’ response) for each individual facial movement pattern in Figure 2-4a. Specifically, each color-coded shape represents an individual observer dynamic facial expression pattern with its hit rate (y axis) and false alarm rate (x axis). Magenta represents Western facial expressions, green East Asian models; triangles represent pain, squares pleasure (see legend in bottom left). The four quadrants that delimited by dashed lines indicate the different response types elicited by the facial expression patterns. For example, high hit and low false alarm rates (top left quadrant) indicate high perceptual discrimination of pain and pleasure, whereas low hits and high false alarm rates (bottom right quadrant) indicate that pain tends to be miscategorized as pleasure (or vice versa). The diagonal dashed line represents an equal rate of hits and false alarms. As shown by the distribution of the data points primarily in the top left quadrant, the vast majority of facial expression patterns of pain and pleasure in both cultures are discriminated well with virtually no instances of ambiguity or confusion (i.e., consistently matching a pain facial expression with a pleasure word). In fact, Figure 2-4a shows that only one of 160 facial movement patterns is ‘ambiguous’ (i.e., in the top right quadrant).

2.3.1.3.1 Physical Distinctiveness Between Facial Expressions of Pain and Pleasure – Bayesian Classification

Perceptual distinctiveness should rely on physically distinct patterns (Bradbury & Vehrencamp, 1998). Therefore, the diametrically opposite messages of pain and pleasure should be communicated using distinctive facial expression patterns. To objectively examine whether facial expressions of pain and pleasure are physically distinct, Classification performance, I used a Bayesian classification approach using a split-half method (1000 iterations) applied to the binary AU patterns. Specifically, I trained the Bayesian classifier using half of facial expression models of pain and pleasure randomly sampled without replacement and tested classification performance using the remaining half of the facial expressions models. Figure 2-4b shows the results. Each color-coded matrix shows the average probability (computed across 1000 iterations) of a pain or pleasure classification for each facial expression type for each culture separately. As shown by the dark-red diagonal squares (correct classifications) in each matrix and corresponding blue off-diagonal squares (confusions), facial expression patterns of pain and pleasure are clearly discriminated (i.e., average classification performance: Pain, Western, $M = 96.95\%$, $SD = 0.44\%$, East Asian, $M = 99.98\%$, $SD = 0.01\%$; Pleasure, Western = 98.61%, $SD = 0.17\%$, East Asian, $M = 95.79\%$, $SD = 0.54\%$).

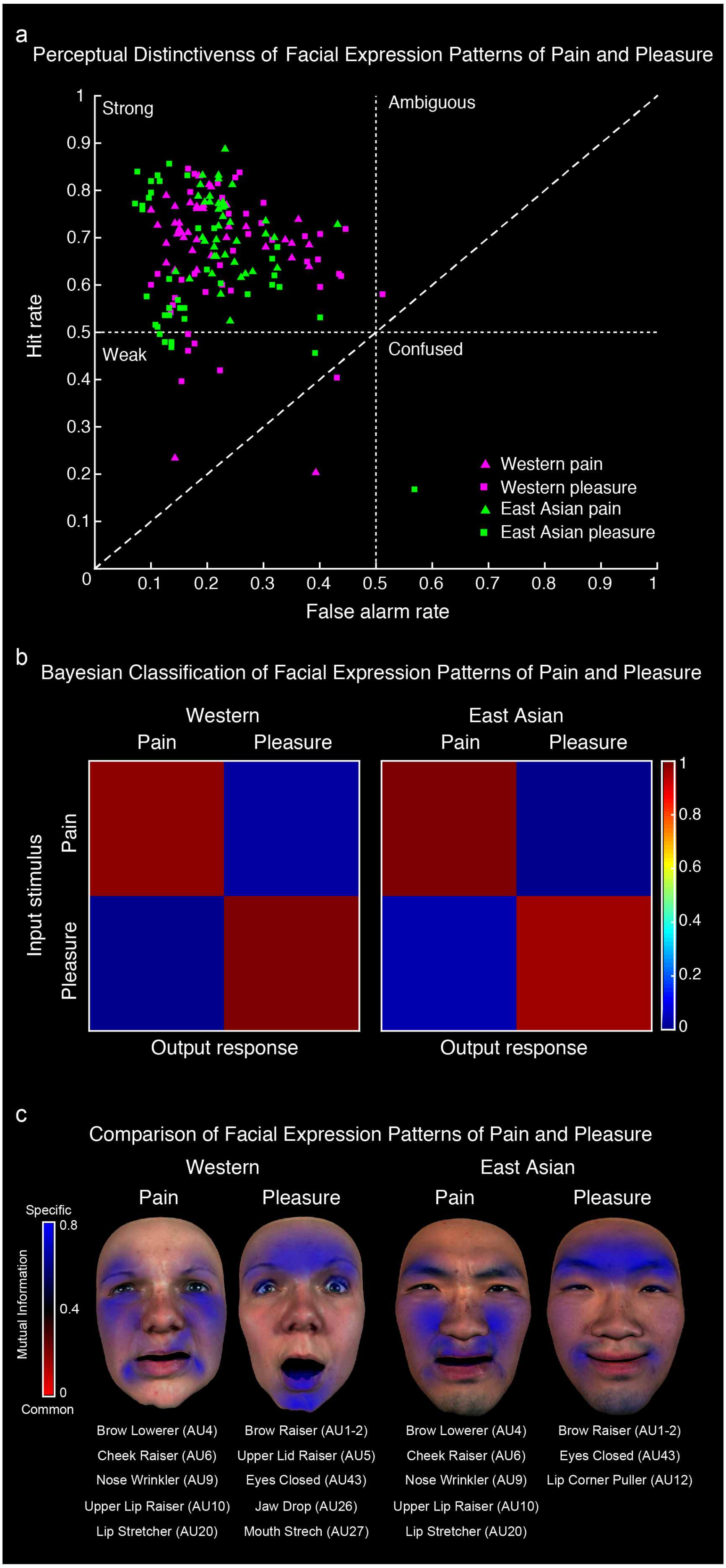

Figure 2-4. Perceptual and physical discrimination of facial expression patterns of pain and pleasure in Western and East Asian culture. (a) *Perceptual discrimination of dynamic facial expression patterns of*

pain and pleasure. Each color-coded shape represents an individual observer facial expression pattern plotted according to its hit and false alarm rate (i.e., d-prime value). Magenta points represent Western facial expressions, green represents East Asian facial expressions; triangles represent pain, squares represent pleasure (see legend in bottom left). The four quadrants (delimited by dashed lines) indicate the different response types elicited by the facial expression patterns (see labels in each quadrant). For example, a high hit and low false alarm rate (top left quadrant) indicates strong perceptual discrimination of pain and pleasure, whereas a low hit and high false alarm rate (bottom right quadrant) indicates that the facial expression tends to be miscategorized. The diagonal dashed line represents an equal rate of hits and false alarms. As shown by the distribution of the data points, the vast majority of facial expression patterns of pain and pleasure in both cultures are discriminated well with virtually no instances of ambiguity (top right quadrant) or confusion (bottom right quadrant). (b) *Bayesian classification performance of facial expression patterns of pain and pleasure*. To objectively examine the physical distinctiveness of the facial expression patterns of pain and pleasure, I applied a Bayesian classifier to the AU patterns of pain and pleasure in each culture separately using a split-half method. Color-coded matrices show for each culture the average probability of a pain or pleasure classification (Output response) for each facial expression type (Input stimulus). Dark red squares show high probability of classification; blue indicates low probability (see colorbar to right). As shown for each culture, facial expressions of pain and pleasure are discriminated with high accuracy. (c) *Comparison of facial expression patterns of pain and pleasure*. To objectively identify the face movements (i.e., Action Units; AUs) that are specific to or common across facial expressions of pain and pleasure in each culture, I used Mutual Information (MI; see *Comparison of Facial Expression Patterns of Pain and Pleasure*). In each culture, color-coded face maps show the AUs that are specific to pain and pleasure (blue coloring, significantly high MI, $p < 0.05$) and those that are common across pain and pleasure (red coloring, low MI; see colorbar on left). In each culture, homogenous blue coloring shows that the facial expression patterns of pain and pleasure are highly distinct with no common AUs. AUs listed below each face map are those specific to each facial expression type.

2.3.1.3.2 Comparison of Facial Movement Patterns of Pain and Pleasure.

The results of the human and Bayesian discrimination task both suggest that facial expressions of pain and pleasure should each comprise a distinct set of face movements. To identify the specific facial movements (i.e., AUs) that are common and those that are distinct between physical pain and pleasure, I applied a two-step analysis where I first for pain and pleasure separately the highly frequent AUs using a Monte Carlo simulation method (*Step 1: Statistically Highly Frequently AUs)* and then used an information theoretic analysis called Mutual Information (MI). MI measures in bits the relationship between the type of messages (i.e., pain or pleasure) and the presence/absence of each AU (e.g., Brow Lowerer) in each culture. Therefore, low MI indicates that the pain and pleasure share that particular AU in their facial expression models (i.e., common AUs) whereas high MI indicates that the AU is mostly present in the models of either pain or pleasure (i.e., distinct AUs).

Step 1: Statistically Highly Frequent AUs. For each culture separately, I first identified the AUs that are statistically highly frequent in the facial expressions of pain or pleasure using a Monte Carlo simulation method. To illustrate, consider I aim to identify the AUs that are statistically highly frequent across the Western facial expression patterns of pain. First, in the total set of 42 AU × 40 Western facial expressions of pain, I computed the frequency of each AU across all of these 40 facial expressions. For example, Brow Lowerer (AU4) is present a total of 32 times over the 40 facial expression patterns, resulting in a 32/40 frequency. I also computed the total number of significant AUs across the facial expression patterns. For example, there are a total of 204 significant AUs across the 40 Western facial expressions of pain. Next, to determine the statistical significance of each AU frequency, I used a Monte Carlo simulation method to randomly distribute (with replacement) the total number of significant AUs (here, 204 AUs) over the space of 42 AU x 40 facial expression patterns, and compute the resulting frequency of each AU. For example, on this iteration, Brow Lowerer (AU4) might have a frequency of 3/40. Over 1000 iterations, I therefore derived a distribution of AU frequencies, which I used to test the null hypothesis that the observed 32/40 proportion of Brow Lowerer (AU4) in the Western facial expressions of pain is significantly higher than chance — i.e. above the 95th percentile of the randomly generated distribution of AU frequencies (i.e., one-tailed $p < 0.05$). If so, I call Brow Lowerer (AU4) a *highly frequent* AU. I repeated this procedure for each AU and for pain and pleasure in each culture separately.

Step 2: Comparison of Highly Frequent AU Between Pain and Pleasure. Having determined the highly frequent AUs for each conversational message category and culture in Step 1, I then examined the relationship of each AU to culture – i.e., do the highly frequent AUs distribute evenly across the facial expressions of both pain and pleasure, or are they highly distinct? To address this question, I quantified the relationship between each highly frequent AU and the type of facial expression (i.e., pain or pleasure) using Mutual Information (MI). Low MI indicates that pain and pleasure share that particular AU across their facial expression patterns whereas high MI indicates that the AU is mostly present in only one type of the facial expression. To establish statistical significance, I used a Monte Carlo approach. For each highly frequent AU and social message (e.g., 'pain'), I produced a random distribution of MI values by randomly shuffling the social message of each individual facial expression pattern 1000 times, computing MI for each AU at each iteration, and then taking the maximum MI value across all AUs. I then used the distribution of maximum MI values to identify the AUs with a MI value in the 95th percentile of the distribution (Nichols & Holmes, 2002).

Figure 2-4c shows the results. As predicted by the discrimination tasks, I found no common face movements between pain and pleasure in either culture. Rather, facial expressions of pain and pleasure are each characterized by a distinct set of AUs. Specifically, Westerners use Brow Lower (AU4), Cheek Raiser (AU6), Nose Wrinkler

(AU9), Upper Lip Raiser (AU10) and Lip Stretcher (AU20) for pain and Brow Raiser (AU1-2), Upper Lid Raiser (AU5), Eyes Closed (AU43), Jaw Drop (AU26), Mouth Stretch (AU27) for pleasure. East Asians use Brow Lower (AU4), Cheek Raiser (AU6), Nose Wrinkler (AU9), Upper Lip Raiser (AU10) for pain and Brow Raiser (AU1-2), Eye Closed (AU43), Lip Corner Puller (AU12) for pleasure.

2.3.1.3.3 Cross-Cultural Comparison of Facial Expression Patterns of Pain and Pleasure

It has been argued that the primary function of facial expressions of physical pain and pleasure is to benefit the expresser by regulating sensory experience – for example, to get rid of painful stimulus by closing the eyes or suppressing the pain sensation by compressing the lips (Frijda, 2002). Others argue that facial expressions of pain and pleasure are used for social communication purposes as they provide information about the subject’s physical states to onlookers for their assistance (e.g., Prkachin, Berzins, & Mercer, 1994; Williams, 2002) or reward (e.g., Fernández-Dols et al., 2011). In comparison to facial expression patterns across cultures can provide insights into whether they serve a communicatory function since culture may have tuned the facial expression pattern depending on the socio-cultural niche (Shariff & Tracy, 2011). In contrast, if face movement patterns primarily serve a physical purpose, I might expect them to be similar across cultures.

To compare the facial expression patterns of pain and pleasure across cultures, I applied the same procedure of analysis (i.e., highly frequent AUs + MI) to quantify the relationship between each highly frequent AU and culture. Specifically, I examined the relationship of each AU to Western and East Asian – i.e., do the highly frequent AUs of pain and pleasure distribute evenly across cultures, or are they specific to only one? To address this question, I quantified the relationship between each highly frequent AU and each culture using Mutual Information (MI). High MI indicates a strong relationship between an AU and the culture – i.e., the AU is largely specific to Western or East Asian – whereas low MI values indicate a weak relationship – i.e., the AU is common across culture. To establish statistical significance, I used the Monte Carlo approach as above. For each highly frequent AU in Western or East Asian culture, I produced a random distribution of MI values by randomly shuffling the culture assignment (i.e., either Western or East Asian) of each individual facial expression pattern 1000 times, computing MI for each AU at each iteration, and then taking the maximum MI value across all AUs. I then used the distribution of maximum MI values to identify the AUs with a MI value in the 95th percentile of the distribution (Nichols & Holmes, 2002).

Figure 2-5 shows the results of cross-cultural comparison. Here, low MI values indicate the AUs that are common across cultures, whereas high MI values indicate the AUs that are specific to one culture (see colorbar to right). Results show that observers from both cultures use a same set of AUs for pain (i.e., Brow Lowerer (AU4), Cheek Raiser (AU6), Nose Wrinkler (AU9), Upper Lip Raiser (AU10) and Lip Stretcher (AU20)). In contrast, both cultures use Brow Raiser (AU1-2) and Eyes Closed (AU43) for pleasure but there are cultural differences. Specifically, Western observers use Upper Lid Raiser (AU5), Jaw Drop (AU26) and Mouth Stretch (AU27), whereas East Asian observers use Lip Corner Puller (AU12).

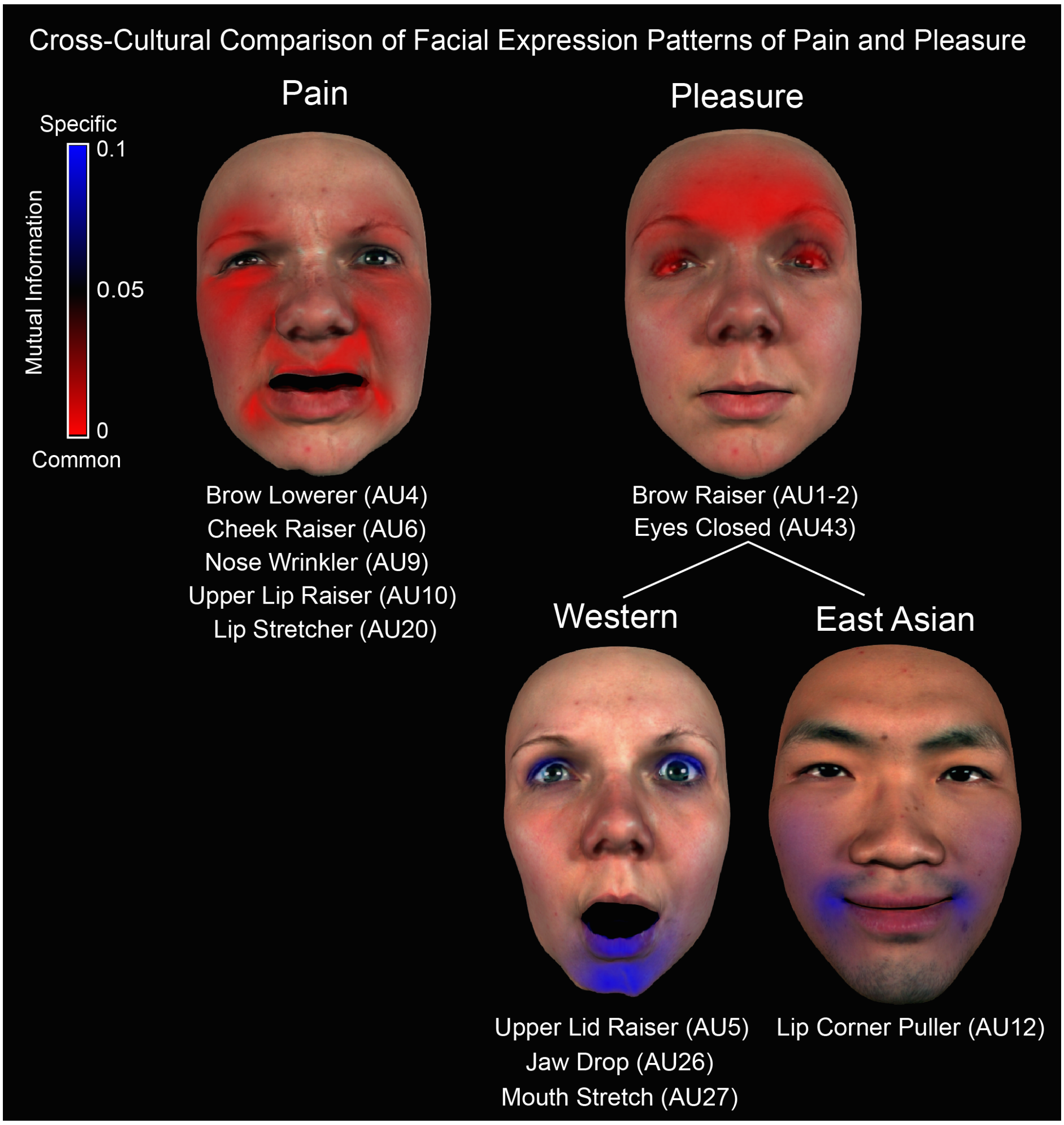

Figure 2-5. Culturally common and culture-specific Action Units in the facial expression patterns of pain and pleasure. To identify in the facial expression patterns of pain and pleasure any AUs that are

culturally common and those that are culture-specific, we used MI (see full details in *Supplemental Materials, Cross-cultural Comparison of Facial Expression Patterns of Pain and Pleasure*). Each color-coded face map shows the AUs that are common across cultures (red, significantly low MI, $P < 0.05$) or specific to one culture (blue, high MI, see colorbar to left). As shown by the red coloring, the facial expressions of pain showed many culturally common AUs (see AU labels below) and no culture-specific AUs. Facial expression patterns of pleasure also showed culturally common AUs (e.g., Brow Raiser – AU1-2) but also culture-specific accents such as Jaw Drop (AU26) and Mouth Stretch (AU27) amongst Western patterns, and Lip Corner Puller (AU12) amongst East Asian patterns.

2.3.1.4 Discussion

In this study, I mathematically modelled the dynamic facial expressions of pain and pleasure in two cultures using reverse correlation and a dynamic facial expression generator. Analysis of the resulting models showed that facial expressions of pain and pleasure are highly distinct, which is demonstrated by both human observers and Bayesian classification performance. Specifically, MI analysis confirmed that a distinctive set of facial movements are used for communicating pain and pleasure in Western and East Asian culture. Furthermore, cross-cultural analysis showed that facial expressions of pain are similar across cultures on the basis of shared face movements including brow lower/frown, eye contraction, wrinkled nose and a horizontally stretched mouth. In contrast, facial expressions of physical pleasure (i.e., pleasure) showed clearly cultural specificities – Western models showed wide opened eyes and a vertically stretched mouth, whereas East Asian models included smiling.

Highly Distinct and Diagnostic Facial Expressions of Pain and Pleasure

In both cultures, I show that facial expressions of pain and pleasure are both *physically* and *perceptually* distinct, thereby challenging current notions that extreme positive and negative affect cannot be distinguished from the face (e.g., Aviezer et al., 2015; Aviezer et al., 2012; Barrett et al., 2011). My data support the views of facial expressions as efficient social signals for communicating valence by showing that pain and pleasure have minimal overlap (i.e., distinct facial movement patterns) to optimize the transmission of valence information (e.g., Darwin, 1872/1998; Smith et al., 2005). My results therefore align with previous theories/studies on facial expressions of emotions that support the distinct facial expressions for pain and pleasure. Specifically, researchers argue that humans interpret their internal experiences and corresponding facial expressions along a continuum of displeasure and pleasure – i.e., valence (e.g., Osgood, 1966; Russell, 1980; Watson & Tellegen, 1985). Such a distinction reflects cross-linguistic evidence that the concepts ‘good’ and ‘bad’ are considered independent semantic primitives (e.g., Wierzbicka, 1999), and that facial expressions are perceived along a continuum of valence during early childhood (e.g., Widen, 2013; Widen & Russell, 2008). Further, theories of facial expressions of emotions argue that facial movement patterns should become more distinct as the internal states become more intense (e.g., Calder et al., 2000; Carroll & Russell, 1996; Hess, Blairy, & Kleck, 1997; Russell, 1980; Tracy, 2014).

Common Biological and/or Culture Roots of Facial Expressions of Pain and Pleasure

My results show the consistency of facial expressions of pain across cultures, supporting views of their culturally common physiological roots to protect from danger and promote recovery (e.g., Darwin, 1872/1998; Grunau & Craig, 1987; Prkachin, 1992; Williams, 2002). Specifically, my data show that pain comprises culturally common face movements that also modulate sensory exposure (Susskind et al., 2008). Specifically, both Westerns and East Asians associated pain with the facial movements enhancing vigilance (e.g., Davis & Whalen, 2001) and sensory rejection (e.g., Chapman et al., 2009; Rozin & Fallon, 1987) – Brow Lowerer (AU4), Nose Wrinkler (AU9) and Upper Lip Raiser (AU10). In contrast, facial expressions of pleasure form distinct clusters within each culture, suggesting the contribution of cultural learning and their role as evolved social signals (Baumeister & Bratslavsky, 1999; Eibl-Eibesfeldt, 1989). Specifically, Westerners associated pleasure with facial movements (i.e., wide opened eyes and a vertically stretched mouth) that accompanying with highly arousal, desirous and intense feeling (Fernández-Dols et al., 2011; Hughes & Nicholson, 2008; Masters & Masters, 1986), which also commonly produced by the excited screaming and loud laugh in human infants (e.g., Darwin, 1872/1998; Ekman, 2006). In contrast, East Asian associated pleasure with much peaceful smiling, suggesting their preference of lower arousal experiences (e.g., Lim, 2016; Tsai, Knutson, & Fung, 2006).

In Study 1, I have examined the role of social communication of facial movement patterns using the diametrically opposite messages of extreme pain and pleasure. Next, I expanded my research to a much broader set of social messages – that is, facial expressions for communicating mental states including those to manipulate the flow of conversation.

3 Study 2: Mapping Dynamic Facial Expressions of A Broad Set of Mental States Across Cultures

3.1 Introduction

To understand how facial expressions are used to communicate mental states in a given culture, it is useful to first examine how cultural individuals conceptualize and organize these mental state terms by examine a core set of mental state terms in each culture. The core set of mental state terms represents the knowledge of the physical and cultural environment (i.e., the external reality) and the understanding of the world (i.e., the internal reality). For example, linguists show that the Sami people who live in the Arctic area have a large number of words for difference types of reindeer (Magga, 2006), whereas the Bedouin people who live in the desert have many different words for camel (Trudgill, 2011). In the similar vein, cultural groups who emphasize the social hierarchy such as Japanese and Korean have developed a sophisticated semantic network of honorifics, whereas other languages (e.g., English) do not. Thus, examining the semantic meaning of such mental state terms (e.g., the level of valence and arousal, Russell, 1980) and the similarities/differences between these terms can provide an substantial understanding of how cultural individuals perceive such social messages and use this knowledge in daily communication (e.g., Nuyts & Pederson, 1999; Sowa, 2014).

Facial expressions can be used to communicate such mental states. In fact, research have suggested that the most pervasive forms of facial expressions in real social contexts is for communicating a variety of mental states (e.g., Rozin & Cohen, 2003). Yet, our knowledge of facial expressions is much limited by the terms of emotions, whereas previous psycho-lexical research suggested that mental state terms is a much broader set of words including belief, intentions and desires (e.g., Allport & Odbert, 1936). Thus, it remains unclear that whether and how facial movement patterns can be used for communicating a broad set of mental states.

This gap of knowledge then limits our view of understanding the variety of facial expressions used in our daily life. Firstly, a central debate in the research of facial expressions of emotions is whether a set of facial movement patterns can be used to communicate one (e.g., Ekman, 1993; Etcoff & Magee, 1992) or more than one meaning (e.g., Aviezer et al., 2012; Carroll & Russell, 1996; Russell, 1997). Understanding how facial movement patterns are used to communicate a broad set of mental states can provide a more comprehensive perspective to address this debate by expanding our knowledge about the communication functions of facial expressions – for example, whether mental states with similar meaning was communicated using similar facial movement patterns and whether a given set of facial expressions can elicit the detection of one or more such social messages.

Secondly, although recent research on the facial expressions of emotions have made significant improvement for our knowledge for facial expressions across cultures (Barrett & Gendron, 2016; Ekman, 2006; Elfenbein, 2013; Jack, Garrod, et al., 2012; Wallraven et al., 2015), it still remains unknown whether there is a set of culturally common facial expressions that are used to communicate the variety of mental states. For example, Jack et al. (2016) modeled over 60 emotions in Western and East Asian culture and suggested that there are four facial movement patterns that are culturally common – it should be noted that facial expressions used in our daily life can delivery a much boarder set of social messages than emotions. More importantly, as suggested by the evaluation theory of facial expressions (Shariff & Tracy, 2011), a set of fundamental facial expressions may have developed from an early stage as sensory regulators (Chapman et al., 2009; Lee et al., 2014) and remain their movement patterns as their physiological originals (Susskind et al., 2008), whereas other facial movement patterns – for example, conversational facial expressions which primarily delivery the message of cognitive states, may have evolved later for the function of social communication and therefore may acquire more complex facial movement patterns by either combining the fundamental facial expression patterns or developing other specific facial movement patterns. Since these complex facial expressions may have developed at later stage and their main function is social communication, it is possible that cultural context played an important for tuning such facial movement patterns. Expanding our knowledge of facial expressions to a broad set of mental states and cultures then cast insight into the understanding of functions of facial expressions and the structure of facial movement patterns as a powerful communication system in social communication.

Finally, with the recent rise of the digital economy, the rapid development of in human-computer interaction requires the knowledge of a broad set of facial expressions in designing and developing socially interactive conversational agents (e.g., Cassell, 2000; De Rosis, Pelachaud, Poggi, Carofiglio, & De Carolis, 2003; Pelachaud, Badler, & Steedman, 1996; Poggi & Pelachaud, 2000). To reliably impact human social judgment, digital agents must be socially aware and generate realistic communication behaviours including facial expressions. To be able to generate a broad set of facial expressions, computer scientists have developed mathematical models to generate facial movement patterns using the two-dimensional (i.e., valence – arousal) or three-dimensional (i.e., valence – arousal – dominance) descriptors of semantic expressivity (e.g., Boukricha, Wachsmuth, Hofstätter, & Grammer, 2009; Jia, Wu, Zhang, Meng, & Cai, 2014; Zhang, Wu, Meng, & Cai, 2007). The prerequisite of such generative system to generate natural and realistic facial expressions, however, is the knowledge of how facial expressions are used in human-human communication. Thus, the knowledge gap of a broad set of facial expressions then prevents developing such social interactive agents and its application with the other sources of multi-modal information such as vocalizations, averted gaze, or hand gestures (e.g., Chovil, 1991; Harness Goodwin & Goodwin, 1986; Mahmoud & Robinson, 2011).

To understand the facial expressions used for communicating a broad set of mental states, I first extract the core set of mental state terms in Western and East Asian culture (i.e., *Experiment 1*) and then modelled the facial movement pattern for each mental state (i.e., *Experiment 2*). To ensure each facial expression pattern communicates the corresponding mental state, I validated each facial expression pattern with a new group of participants in each culture using the approach of signal detection theory (i.e., *Experiment 3*).

3.2 Experiment 1: Mapping the Semantic Network of Mental States Across Cultures

To extract mental state terms in Western and East Asian culture, I first derived a core set of highly familiar and highly typical mental state terms. To examine the semantic network of these mental state terms, I then measured the meaning of each mental state category on three dimensions – valence, arousal and dominance. To illustrate the methods I used to extract the mental state terms, I will now present a review on the previous work of understanding the semantic network of mental states in Western (English) and East Asian (Chinese) culture.

3.2.1 Mental State Terms in Western Culture (English language)

Although the mental state terms were discussed as the representation of the knowledge of perception and cognition from a long time (e.g., Aristotle, 1885/1984; Descartes & De Spinoza, 1619/1961), the systematic study of extracting and examining a set of words in the English language that describing ‘mental states’ began with Allport and Odbert (Allport & Odbert, 1936). Initially, the main aim of this study was to provide a psycho-lexical classification of personality-descriptive terms (i.e., ‘trait-names’), which later became the major foundation of the lexical hypothesis in the psychological research of personality. To provide an exhaustive list of personality-descriptive terms, the authors extracted the ‘trait-names’ from the second edition Webster’s Unabridged Dictionary of the English Language (1925). However, during this process, the authors found that there is a large set of words is used for describing the temporary forms of internal states. The researchers finally decided to provide an individual list (i.e., Column II) for these ‘mental state terms’ using the definition of “present activity, temporary states of mind, and mood.” As a result, Allport and Odbert (1936) extracted a total of 4541 terms of mental states from the second edition of Webster’s Unabridged Dictionary of the English Language (1925).

Although the authors provided a relatively complete list of mental state terms, it is not clear that whether these words were frequently used by the population in terms of their familiarity and their meaning that used for describing mental states (i.e., typicality). For the purpose of providing an exhaustive list of such terms, their methods were relatively liberal and subjective. It should be noted that this study was done nearly a hundred years ago. English vocabulary and the way people use a particular term have changed dramatically since then. For example, Allport and Odbert’s list included the words such as *cross*, *eaten* and *wood*, which are not typically categorized as mental states by the native speakers nowadays.

Mental state terms are typically used in modern psychology to measure the ‘theory of mind’ (Premack & Woodruff, 1978). For example, the use of mental state terms in children’s conversations reflects how children acquire the knowledge of beliefs, intentions and desires of others (e.g., Johnson & Wellman, 1980; Moore, Bryant, & Furrow, 1989; Shatz, Wellman, & Silber, 1983). However, the set of mental state terms used in these development psychology studies is usually very small (e.g., two or three mental state terms such as *know* and *guess*). The criterion of the word selection is often based on common sense and therefore can be arbitral. Baron-Cohen et al. (1994) measured the ‘theory of mind’ in the children with Autism Spectrum Disorder using a list of 40 mental state terms, but it is not clear how and why these mental state terms but not other words were selected. Apart from their role of measuring internal states in psychological research, the special function of mental state terms in conceptualization has aroused the interest of linguists (e.g., Croft, 1993). For example, d'Andrade (1987) proposed ‘a folk model of mind’ where he suggested there are six terms of mental state terms in English – *perception*, *belief/knowledge*, *feeling/emotions*, *desires/wishes*, *intentions* and *will* – and provided a list of words which belongs to each of these terms. In summary, previous studies in psychology and linguistics on mental state terms provide a rich source for understanding their role in social communication. Yet, most studies used the mental state terms in a more descriptive and subjective way. It remains unknown what is the core set of mental state terms (i.e., highly frequent and highly typical in terms of their meaning as describing the internal states) used in English language and how they are perceived by the population.

3.2.2 Mental State Terms in East Asian Culture (Chinese language)

Research of mental state terms in Chinese language mental states description began with Mashi Wentong 《马氏文通》(Ma's Grammar; Ma, 1898). This is the first systematic work for explaining Chinese Grammar and how different types of words are used to construct sentences in Chinese. Specifically, the author suggested that the words that describe the mental activity (心), feeling (感) and thoughts (意) are 'implicit terms' since they are not as 'visible' as the other terms that describe the human behaviour (e.g., eat, drink). The authors concluded that these mental state terms are semantically different and their functions in Chinese language and Chinese philosophy of mind should be further examined. It should be noted that the modern Chinese language system replaced the pre-modern Chinese language as the official language of China around 1950s (Chen, 1999). For the aim of my thesis, I will only focus on the mental state terms in the modern Chinese language. To taxonomize and understand the functions of mental state terms, linguists extracted mental state terms from Dictionary of Chinese Terms (e.g., Wen, 2007)， Dictionary of Using Chinese Verbs (e.g., Cang, 2012; Ding, 2008; Sun, 2008; Zhu, 2012) and the Corpus for Chinese Linguistics (Li & Zhou, 2015; Qiao, 2015). Wen (2007) suggested that there are three main types of mental state terms in Chinese languages – emotions (e.g., happy 喜, anger 怒, sad 哀), perception/cognition (e.g., thinking 思, learning 学, remembering 记) and intentions (e.g., planning 打算). To examine the of the core set of mental state terms, Zhang and Lu (2007) extracted a set of 80 core mental state terms from 18 key literature sources on Chinese mental state terms (Hao, 1999; Jingyu, 2004; Wang, 2004; Yang, 1999). Although these studies provide a rich source for studying a broad range of mental state terms in Chinese language, the criteria for selecting these mental state terms is relatively descriptive and subjective. It is not clear whether these mental state terms can reflect the core set of mental state terms (i.e., highly familiar and typical for describing mental state) used by a large set of native Chinese speakers.

3.2.3 Current Work

Based on the previous studies on mental state terms in English and Chinese, I aim to first provide an exhaustive list of mental state terms in both languages by combining the existing lists with the mental terms used in key literature sources (i.e., Step 1). To extract the core set of mental states in English and Chinese, I then measured their familiarity (i.e., Step 2) and prototypicality (i.e., Step 3) using ratings from native speakers. Finally, to examine the semantic similarity/differences of these mental state terms, I measured the level of valance, arousal and dominance of each mental state term (i.e., Step 4).

3.2.3.1 Step 1: Extract Terms From Key Literature Sources.

I first provided an exhaustive list of mental state terms in English by extracting the 4549 mental state terms from Allport & Odbert's list of 'temporary states of mind, and mood' (page 26, Column II, Allport & Odbert, 1936) and combining with a total of 325 mental state terms extracted from key literature sources (Back et al., 2009; Baron-Cohen, 2003; Baron-Cohen et al., 1996; Baron-Cohen et al., 2001; Baron-Cohen, Wheelwright, Jolliffe, & Therese, 1997; Bavelas & Chovil, 2000; Bavelas et al., 2000; Bretherton & Beeghly, 1982; Brown, Donelan‐McCall, & Dunn, 1996; el Kaliouby, 2005; El Kaliouby & Robinson, 2005; Furrow, Moore, Davidge, & Chiasson, 1992; Golan, Baron-Cohen, & Golan, 2008; Golan, Baron-Cohen, & Hill, 2006; Lee, Harkness, Sabbagh, & Jacobson, 2005; Mátyássy, Kelemen, Sárközi, Janka, & Kéri, 2006; Nusseck, Cunningham, Wallraven, & Bülthoff, 2008; Shatz et al., 1983; Symons, 2004; Tager-Flusberg & Sullivan, 1995; Tager‐Flusberg, 1992; Taumoepeau & Ruffman, 2006; Zeng, Pantic, Roisman, & Huang, 2009). I then removed all repeated words, resulting in a total of 4784 mental state terms.

I applied the same methods for the list of Chinese mental state terms by extracting 765 Chinese mental states from Wen (2007) using the definition of 'describing mental activity (心理活动) or mental states (心理状态) such as emotion (感情), intention (意向) and cognition (认知)' and combining with a total of 775 mental state terms extracted from key literature sources (Cang, 2012; Chen, 2009; Chen, 2008; Ding, 2008; Feng, 2014; Han, 2006; Hao, 1999; Hao, 2011; Jiang, 2004; Jinjin, 2013; Lan, 2014; Li & Zhou, 2015; Li, 2013; Qi, 2007; Run, Zhang, Zhang, & Sun, 2012; Sun, 2008; Wang, 2002, 2004; Wang, 2012; Xu & Wang, 2005; Yang, 1999; Zhang, 2012; Zhang & Liu, 2011; Zhou & Shao, 1993; Zhu, 2012). I then removed all repeated words, resulting in a total of 824 Chinese mental state terms. It should be noted that the Chinese mental state terms are much fewer than the English ones because China developed a completely new system of their language (i.e., Modern simplified Chinese), which only keeps the modern Chinese terms in the Dictionary of Modern Chinese (1978).

3.2.3.2 Step 2: Remove Quirky and Old Fashioned Terms.

Next, to ensure that all terms are used frequently in contemporary discourse, I removed all terms with low frequency (i.e., below 10 per million for both English and Chinese) in the Corpus – the British National Corpus and the Modern Chinese Corpus. As a result, I obtained a total of 2077 English mental state terms and 614 Chinese terms. In addition, for both languages, I asked native speakers to remove any terms considered quirky, old fashioned or rarely used in contemporary discourse as follows.

3.2.3.2.1 Participants

I recruited 16 native English speakers (16 European; 8 male; mean age 24 years, SD 3.5 years) and 16 native Chinese speakers (16 Chinese; 8 male; mean age 23 years, SD 2.3 years). All participants had normal or corrected-to-normal vision, were free from any lexical, reading, language or visual condition affecting the processing of terms (e.g., dyslexia) as per self-report. I paid each participant £6 per hour, and obtained their written informed consent. I used the same criteria for the recruitment and payment in all subsequent experiments. The University of Glasgow College of Science and Engineering Ethics Committee authorized the experimental protocol (Reference No: 300140082).

3.2.3.2.2 Procedure

Participants received an online list of all mental state terms in their native language. The participants were instructed to “highlight all terms they considered quirky, old fashioned or rarely used in contemporary discourse”. The order of mental state terms is random for each participant. All terms are present using lower case font (‘MS Sans Serif’ for English and ‘MS Song (宋体)’ for Chinese) and in Font size 16. Participants have no time limitation to finish their task.

3.2.3.2.3 Results

From the mental state word lists, I removed 306 English terms and 5 Chinese terms indicated as "quirky, old fashioned or rarely used in contemporary discourse" by a large majority of the native speakers (i.e., 80%) in each culture. As a result, I obtained a total of 1771 English terms and 609 Chinese terms that are highly frequent in Western and East Asian culture.

3.2.3.3 Step 3: Rate the Typicality of Mental States.

To extract highly typical mental state terms in each culture, I asked a new group of native speakers to rate each word accordingly.

3.2.3.3.1 Participants

I recruited 36 native English speakers (33 European, 3 North American; 18 male; mean age 24 years, SD 5.8 years) and 36 native Chinese speakers (36 Chinese; 18 male; mean age 23 years, SD 2.7 years).

3.2.3.3.2 Procedure

On each experimental trial, participants viewed a word and rated it according to whether it is typical of a mental state on a 7-point Likert scale ('not at all a mental state' to 'definitely a mental state'). In the instructions, I provided the definition of 'mental state' as "states of mind that people exhibit, express and attribute to each other, including cognitive states, intentions, beliefs, desires and focus of attention" in accordance with the previous literature (Allport & Odbert, 1936; Norman, 1967; Premack & Woodruff, 1978, Fan et al. 1999, Wen 2007). Participants had the option to select 'I don't know this word' if they were unfamiliar with or were not sure of the meaning of the word. Participants rated all terms in their native language, which I presented in random order across the experiment. Each word remained visible until response. I presented all terms on a 19- inch flat panel Dell monitor with a refresh rate of 60 Hz and resolution of 1024 × 1280, with each word presented in lower case white font (English: 'MS Sans Serif,' mean size: 1.74 cm [SD 0.32 cm] × 5.67 cm [SD 1.73 cm]; Chinese: 'MS Song,' mean size: 1.59 cm [SD 0.12 cm] × 4.12 cm [SD 1.56 cm]) on a black background in the centre of the screen. A chin rest ensured a constant viewing distance of 80 cm, subtending 1.74° (vertical) and 4.06° (horizontal) of visual angle. The 7-point rating scale appeared beneath each word, with the ascending/descending order of the scale counterbalanced across participants. Participants used a Graphic User Interface (GUI) to submit their response. I used Matlab 2012a to present the stimuli and record responses.

3.2.3.3.3 Results

Based on the ratings obtained in the above task, I identified all terms rated as highly typical (i.e., >=6) by a large majority of the native speakers (i.e., 80%), resulting in a total of 117 English terms and 111 Chinese terms. Finally, I transformed all English terms to adjectives (e.g., ‘doubt’ became ‘doubtful,’ ‘embarrassment’ became ‘embarrassed’) and removed 10 repeated terms. I did not transform the Chinese mental state terms because Chinese do not differ across different word types. In total, using the criteria detailed above I selected a set of 107 English and 111 Chinese core mental states (see Table 3-1 for a full list of English and Chinese mental state terms).

Table 3-1

Core Set of Mental State Terms in English and Chinese

English Mental State Terms		Chinese Mental State Terms (with English Translation)			
afraid	impatient	仰慕	admiration	倾慕	adore
agitated	indecisive	反省	introspection	恨不得	itch to
amazed	infatuated	羡慕	begrudge	恨不能	be vexed at not being able to
amused	infuriated	藐视	contempt	恼	annoyed
angry	insecure	警觉	vigilance	恼怒	exasperation
annoyed	inspired	讨厌	disgusting	悔恨	regret deeply
anxious	interested	贪恋	cling to	悔悟	awake from sin
apprehensive	intimidated	迷恋	infatuated	惊恐	panic
ashamed	intrigued	钟爱	favorite	惧怕	fear
astonished	irritated	钦佩	esteem	想念	miss
astounded	jealous	震怒	wrath, ouraged	愁	anxious
bored	joyful	震惊	shock	兴奋	excited
calm	lonely	发怒	rage, enraged	愉快	delightful
concerned	mad	高兴	happy	感兴趣	interested
confident	mesmerized	鼓励	encourage	感动	moved
confused	miserable	发愁	worried	愤怒	wrath, angry
contented	mortified	发慌	nervous	愤恨	indignantly resent
curious	offended	发毛	scared	憎恨	hostile
cynical	outraged	吃惊	surprised	憎恶	loathe
delighted	overwhelmed	吃醋	jealous	懊悔	regret
depressed	panicked	后怕	fear after the event	担心	afraid
desire	passionate	后悔	repent	担忧	anxious
despair	perplexed	喜欢	like	动心	tempted
disappointed	petrified	伤心	sad	敬仰	venerate
disgusted	pleased	困惑	confused	无聊	bored
disheartened	proud	在乎	care about	暴怒	fury
disinterested	regretful	在意	take notice of	欢喜	joyful
distracted	relaxed	失望	disappointed	沉醉	intoxicated
distraught	relieved	妒忌	jealousness	渴求	be eager for
distressed	sad	嫉妒	jealous	满意	satisfaction
distrustful	satisfied	害怕	afraid	满足	content
disturbed	scared	害羞	shy	激动	excitement
doubtful	shocked	害臊	shy	烦	annoyed
elated	sorry	崇拜	worship	厌恶	disgust
embarrassed	surprised	佩服	have a good opinion of	热爱	enthusiastic
encouraged	sympathetic	崇敬	respect	爱恋	in love with
enraged	terrified	庆幸	rejoice	爱慕	adore
enthralled	thoughtful	开心	happy	牵挂	care
enthusiastic	thrilled	心疼	love dearly	犯愁	worry
excited	tired	忏悔	repent	狐疑	suspicious
exhausted	troubled	忧心	worry	生气	angry
fascinated	unashamed	忧虑	concerned	畏	fear
fearful	uncertain	快乐	happy	畏惧	afraid
frenzied	uncomfortable	怀念	cherish the memory of	疑惧	suspicious and fear
frightened	unenthusiastic	怀疑	doubt	厌烦	bored
frustrated	unimpressed	倾心	fall in love with	疑虑	misgivings
furious	uninterested	怒	rage, enraged	痛惜	deplore
glad	unsure	怕	fear	盛怒	rage, furious
grateful	unsympathetic	怜悯	love tenderly	盼望	look forward to
guilty	upset	恋爱	love tenderly	看不起	despise
happy	worried	思乡	homesick	看得起	think highly of
homesick		思念	miss	着急	vexed
hopeful		思考	thinking, thoughtful	着迷	fascinated
horrified		怨	blame	绝望	despair
hostile		怨恨	resentment	羞	ashamed
humiliated		恨	hate		

3.2.3.4 Step 4: Valence, Arousal and Dominance.

To examine how these core mental state terms are semantically similar and different similarity and differences, I measured each word on three key dimensions (e.g., Warriner, Kuperman, & Brysbaert, 2013) – valence (e.g., happy or unhappy), arousal (e.g., excited or clam) and dominance (e.g., in control or being controlled) using a new set of participants. This three-dimensional model of is often used in the field of computer science for measuring the expressivity of communication behaviors including facial expressions (e.g., Boukricha et al., 2009; Jia et al., 2014; Zhang et al., 2007), which gives insights into my future work of building a generative system of facial expressions for human-computer interactions.

3.2.3.4.1 Participants

I recruited 32 native English speakers (32 European, 16 male; mean age 22 years, SD 2.9 years) and 32 native Chinese speakers (32 Chinese, 16 male; mean age 23 years, SD 1.4 years).

3.2.3.4.2 Procedure

On each experimental trial, participants viewed a word in their native language and rated it according to its perceived valence, arousal and dominance on a 9-point Likert scale from 1 (happy [excited; controlled]) to 9 (unhappy [clam; in control]) depending on the dimension on that trial. Each participant rated all of 107 English terms or 111 Chinese terms that are presented in random order across the experiment. I blocked the dimensions of valence, arousal and dominance and randomized the order of the blocks across participants. Terms remained visible until response. I presented all stimuli on a 19- inch flat panel Dell monitor with a refresh rate of 60 Hz and resolution of 1024 × 1280, with each word presented in lower case white font (English: 'MS Sans Serif,' mean size: 1.46 cm [SD 0.09 cm] × 4.98 cm [SD 1.58 cm]; Chinese: 'MS Song,' mean size: 1.59 cm [SD 0.12 cm] × 4.12 cm [SD 1.56 cm]) on a black background in the centre of the screen. A chin rest ensured a constant viewing distance of 80 cm, subtending 1.05° (vertical) and 3.57° (horizontal) of visual angle. The 9-point rating scale appeared beneath each word and I counterbalanced the ascending/descending order of the scale across participants. Participants used a Graphic User Interface (GUI) to submit their response. I presented all stimuli on a 19- inch flat panel Dell monitor with a refresh rate of 60 Hz and resolution of 1024 × 1280 and used Matlab 2012a to present the stimuli and record responses.

3.2.3.4.3 Results

For each mental state term and culture, I compute the average rating of valence, arousal and dominance across all participants. To examine the main structure of mental state terms, I measure the similarity between each pair of mental sates using 1 – normalized Euclidean distance between their ratings of valence, arousal and dominance. **Figure 3-1** shows the results. For each mental state term and culture, the color-coded circles on the left show the high or low level of valence, arousal and dominance (see the legend above). The satiability of each circle represents the distance towards to the two ends. For example, according to the ratings of valence, high saturation red circles represent positive mental states, whereas high saturation blue circles represent negative mental states. The color-coded matrixes on right show the similarity (i.e., 1 – normalized Euclidean distance) between each pair of mental state terms, where bright yellow represents high similarity and dark blue represents low similarity. **Figure 3-1** shows that there are two main clusters of mental state terms for both cultures, which are mainly based on the level of valence as positive (i.e., the top left cluster) and negative (i.e., the bottom right cluster). Specifically, the number of negative mental state terms is roughly twice of positive mental state terms in both English and Chinese language.

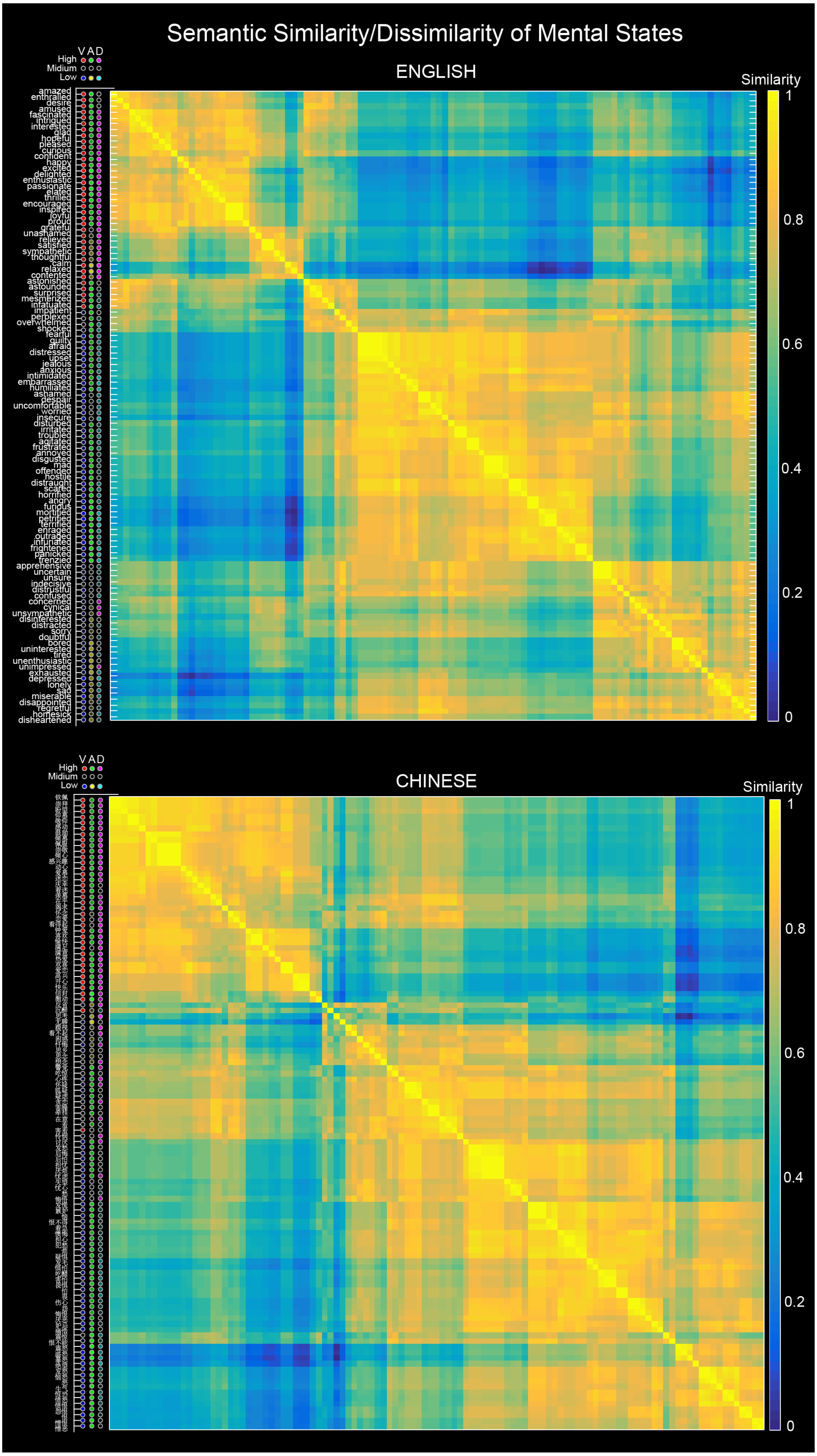
Semantic Similarity/Dissimilarity of Mental States
ENGLISH
V A D
High
Midium
Low
Similarity
1
0.8
0.6
0.4
0.2
0
amazed
enthralled
desire
amused
fascinated
intrigued
interested
glad
hopeful
pleased
curious
confident
happy
excited
delighted
enthusiastic
passionate
elated
thrilled
encouraged
inspired
joyful
proud
grateful
unashamed
relieved
satisfied
sympathetic
thoughtful
calm
relaxed
contented
astonished
astounded
surprised
mesmerized
infatuated
impatient
perplexed
overwhelmed
shocked
fearful
guilty
afraid
distressed
upset
jealous
anxious
intimidated
embarrassed
humiliated
ashamed
despair
uncomfortable
worried
insecure
disturbed
irritated
troubled
agitated
frustrated
annoyed
disgusted
mad
offended
hostile
distraught
scared
horrified
angry
furious
mortified
petrified
terrified
enraged
outraged
infuriated
frightened
panicked
frenzied
apprehensive
uncertain
unsure
indecisive
distrustful
confused
concerned
cynical
unsympathetic
disinterested
distracted
sorry
doubtful
bored
uninterested
tired
unenthusiastic
unimpressed
exhausted
depressed
lonely
sad
miserable
disappointed
regretful
homesick
disheartened
CHINESE
V A D
High
Midium
Low
Similarity
1
0.8
0.6
0.4
0.2
0

Figure 3-1. Similarity matrixes between mental state words based on their level of valence, arousal and dominance. For each culture, the color-coded matrixes show the similarity (i.e., computed as 1 – normalized Euclidean distance) between each pair of mental state terms listed on the left. Bright yellow represents high similarity, whereas dark blue represents low similarity. For each mental state term, the color-coded circles on the left show the high or low level of valence, arousal and dominance (see the legend above). The satiability of each circle represents the distance towards to the two ends. For example, in the ratings of valence, high saturation red circles represent positive mental states, whereas high saturation blue circles represent negative mental states.

3.3 Experiment 2: Modelling Dynamic Facial Expressions of Mental States

For each mental state term and culture, I then mathematically modelled the dynamic facial expression (if there is any facial movement patterns) for communicating that social message. To derive these models, I used a dynamic facial expression generator (Generative Face Grammar, GFG; Yu et al., 2012) combined with the psychophysical method of reverse correlation (Ahumada & Lovell, 1971) and subjective human perception (see also Gill et al., 2014; Jack, Caldara, et al., 2012; Jaworska et al., 2014; Richoz, Jack, Garrod, Schyns, & Caldara, 2015).

3.3.1 Methods

3.3.1.1 Permutation to determine the number of participants and words

Since there is a large set (i.e., > 100) of core mental state terms in each culture, it is not practicable to ask each participant to choose from these words every time. Therefore, for each participant I randomly selected a set of mental state terms with replacement. The results of similarity matrices above show that the main clusters of mental state terms are based on their valence ratings. Thus, to ensure the homogeneity of each set of mental state terms that is present to individual participants, I kept the ratio of positive and negative words as 1/3 and 2/3 for each set of mental state terms.

To determine the optimal number of participants and the number of words presented to each participant, I first run a permutation for each culture separately using the total number of mental state terms (i.e., 107 words for English and 111 words for Chinese) and the number of total participants for each mental state term. For each mental state, I aim to model a set of facial expression patterns based on at least 5 participants in each culture and the total number of participants should be as equal as possible. That is, each mental state should be presented at least 6 times across different participants. For each interaction of the permutation, I randomly chose N different words (from 1 to 20) from the total set of mental state terms with replacement and assign to a set of participants, where each participant see the same number of mental states including 1/3 is positive words and 2/3 is negative words. I keep adding the number of participants until the minimum of the total number of any mental state terms was > 5. For each of N, I did this permutation 1000 times. To ensure that the total frequency of each mental state word is roughly the same, for each of N I only remain the results from these 1000 iterations with the most flat distribution of the frequency of the mental state terms by checking the minimum of differences with the target total frequency of all mental state terms (i.e., Western: 107 words × 6 times; East Asian: 111 words × 6 times). As a result, I obtained for each of N the optimal number of participants I need in each culture. I then ranked the results for different N by their differences with the target total frequency of all mental state terms. Results show that the optimal number of participants is 54 Westerners and 59 East Asians, with each participant see a set of 12 mental state terms – 4 positive words and 8 negative words.

3.3.1.2 Participants

I recruited 54 Western white Caucasian (54 European, all native English speakers, 27 male, mean age 22 years, *SD* = 3.94 years) and 59 East Asian observers (59 Chinese, all native Chinese speakers, 30 male, mean age 24 years, *SD* = 2.30 years). To control for the possibility that any culturally common facial expression patterns could have been learned from cross-cultural interactions, I recruited observers with minimal exposure to and engagement with other cultures (De Leersnyder et al., 2011) as assessed by screening questionnaire (see *Appendix A, Screening Questionnaire*). All East Asian observers had a maximum UK residence of 3 months at the time of testing, and had a minimum International English Testing System (IELTS) score of 6.0 (Competent User). All observers had normal or corrected-to-normal vision and were free from any lexical, reading, language (e.g., dyslexia) or emotion related atypicalities (e.g., Autism Spectrum Disorder, depression, anxiety) as per self-report. I paid each observer £6 per hour, and obtained their written informed consent prior to testing. The University of Glasgow College of Science and Engineering Ethics Committee authorized the experimental protocol (Ref No: 300150031).

3.3.1.3 Stimuli

Figure 3-2 illustrates the stimulus generation and task procedure using one illustrative trial. On each experimental trial, a dynamic facial expression generator – the Generative Face Grammar (GFG; Yu et al., 2012) – randomly selects a biologically legitimate combination of individual facial movements called Action Units (AUs; Ekman

& Friesen, 1978a) from a core set of 42 AUs (n = 5, P = 0.6, median = 3, minimum = 1, maximum = 6). For example, as shown in **Figure 3-2**, on this illustrative trial three AUs are randomly selected – Brow Lowerer (AU4) color-coded in red, Nose Wrinkler (AU9) color-coded in green and Lip Stretcher (AU20) color-coded in blue. For each AU, the GFG then assigns a random movement by selecting random values for each of seven temporal parameters (i.e., onset latency, acceleration, peak amplitude, peak latency, sustain, deceleration, and offset latency; see labels illustrating the blue curve) from a uniform distribution. I generated time courses for each AU using a cubic Hermite spline interpolation (6 control points, 60 time frames). The dynamic AUs are then combined to produce a random facial animation, shown here with four snapshots across time. I displayed each facial animation on one of 10 white Caucasian (10 European, 5 male, mean age 24 years, SD = 5.23 years) or East Asian (10 Chinese, 5 male, mean age 24 years, SD = 1.84 years) identities. I generated all stimuli using a standard procedure in 3D Studio Max (see Yu et al., 2012 for full details).

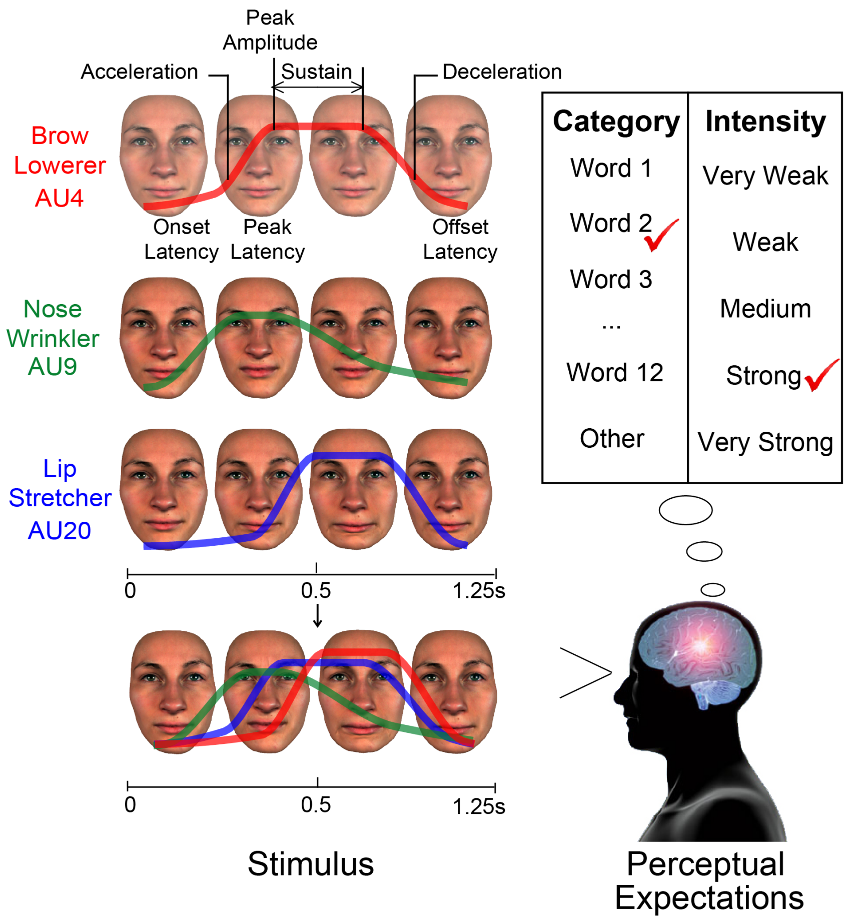

Figure 3-2. Modelling dynamic mental representations of socially interactive facial expressions. Stimulus. On each experimental trial, a dynamic facial expression generator (Generative Face Grammar, GFG; Yu et al., 2012) randomly selects a biologically legitimate combination of individual facial movements called Action Units (AUs; Ekman & Friesen, 1978a) from a core set of 42 AUs (here, Brow Lowerer – AU4 color-coded in red, Nose Wrinkler – AU9 in green, and Lip Stretch – AU20 in blue). The GFG then assigns a random movement to each AU by selecting random values for each of seven temporal parameters (i.e., onset latency, acceleration, peak amplitude, peak latency, sustain, deceleration, and offset latency; see labels

illustrating the blue curve). The dynamic AUs are then combined to produce a random facial animation, shown here with four snapshots across time. The color-coded vector below shows the three AUs randomly selected on this example trial. Mental representations. The observer views the facial animation and if the random movements form a pattern that correlates with their mental representation (i.e., prior knowledge) of one of the mental state terms, they categorize it accordingly and rate its perceived intensity on a 5-point Likert scale. Otherwise, the observer selects 'other.' For each Western and East Asian observer, I pseudo-randomly selected with replacement a set of 12 mental states (i.e., 4 positive mental states, 8 negative mental states) from the pool of 107 English terms or 111 Chinese terms and displayed them in two columns. I presented each observer with the same set of randomly selected terms during the whole experiment. Each observer categorizes 3600 such facial animations displayed on same-race faces.

3.3.1.4 Procedure

Each observer views the facial animation and if the random facial movements form a pattern that correlates with their mental representation (i.e., knowledge) of one of the mental states, they categorize it accordingly and rate the perceived intensity on a 5-point Likert scale. Otherwise, the observer selects 'other.' For each Western observer, I pseudo-randomly selected with replacement a set of 12 English or Chinese mental states from all of mental states (4 positive mental states and 8 negative mental states, which represents the ratio of positive and negative mental states in the whole set of English/Chinese mental states) and displayed them in two columns on the screen. For each observer, the set of mental states are clustered (positive and negative; presented in random order within each cluster) and remain the same during the whole experiment. Each observer (54 Western, 59 East Asian) categorized 3600 such animations displayed on same-race faces. Stimuli (Western Faces: mean size: 19.27 cm [SD 0.83 cm] × 12.38 cm [SD 0.49 cm]; East Asian Faces: mean size: 17.46 cm [SD 0.80 cm] × 10.56 cm [SD 0.35 cm]) appeared in the observer's central visual field and played only once for a duration of 1.25s. A chin rest ensured a constant viewing distance of 78 cm, with images subtending 14.08° (vertical) and 9.08° (horizontal) of visual angle, reflecting the average size of a human face (Ibrahimagić-Šeper et al., 2006) during natural social interaction (Hall, 1966). I presented the response options in the observer's native language (i.e., English or simplified Chinese). I presented each of the mental state terms and the intensity labels (English: mean size: 1.40 cm [SD 0.17 cm]) × 3.75 cm [SD 1.00 cm]. Chinese: 1.20 cm [SD 0.12 cm]) × 3.65 cm [SD 0.54 cm]) in lower case white font ('MS Sans Serif') and subtended 1.03°(vertical) × 2.76° (horizontal) of visual angle.

3.3.1.5 Results

Following the experiment, I used an established reverse correlation analysis to identify the dynamic AUs that are significantly correlated with the perception of each mental state for each individual observer in each culture as follows. First, I performed a Pearson correlation for each individual AU between two binary vectors– the first vector recorded the presence vs. absence of the AU considered on each trial; the second vector recorded the responses of the observer on each corresponding trial. For all AUs significantly correlated with a given response (e.g., 'doubtful'), I assigned a value of 1 (two-tailed $p < .05$) and 0 otherwise, resulting in a 1×42-dimensional binary vector. Thus, each resulting vector details the AUs that are significantly correlated with the perception of each mental state for each individual observer. I did not analyze trials categorized as 'other' because they do not correspond to any specific social message.

For each significant AU in the 1×42-dimensional binary vector, I also computed an estimate of its temporal dynamics as follows. For each of the seven temporal parameters (i.e., onset latency, acceleration, peak amplitude, peak latency, sustain, deceleration, and offset latency, cf. labels illustrating the blue curve in **Figure 3-2**. I performed an independent linear regression between the observer's intensity ratings (e.g., 'very strong') and all trials where the observer selected the mental state in question (e.g., 'doubtful'). I removed the facial expression patterns without any significant AU and obtained a total of 490 Western and 499 East Asian dynamic facial expression patterns. Each facial expression pattern is represented as a 1×42-dimensional binary vector detailing the significant AUs, plus 7 values detailing the temporal parameters of each significant AU. Computing these dynamic facial expression patterns in this way enables their reconstruction as stimuli for subsequent behavioral tasks. To derive movies of the resulting dynamic facial expression patterns to use as stimuli in Experiment 3, and for illustration purposes, I combined the significantly correlated AUs with their corresponding temporal parameters derived from the regression coefficients and displayed the dynamic pattern on different face identities. For stimuli, I used the facial expression patterns derived from 'high intensity' ratings as these represent the most salient signals.

3.4 Experiment 3: Validating Facial Expressions of Mental States

Next, I validated each of the dynamic facial expression patterns by measuring its within-culture communication accuracy using d-prime (Green & Swets, 1966). D-prime provides a reliable measure of communication accuracy as it considers both hit rates (signal present, 'yes' response) and false alarm rates (signal absent, 'yes' response), which safeguards against artificially high accuracy rates (e.g., see Elfenbein, Mandal, Ambady, Harizuka, & Kumar, 2002; Lynn & Barrett, 2014; Russell, 1994). D-prime also provides a measure that reflects key and independent components of communication – signal discriminability (e.g., see Guilford & Dawkins, 1991) and biases in the response characteristics of the observer (e.g., see Stanislaw & Todorov, 1999).

3.4.1 Methods

3.4.1.1 Validators

I recruited a new set of 40 Western white Caucasian validators (40 European, 20 male, mean age 21 years, SD = 3.33 years) and 40 East Asian validators (40 Chinese, X male, mean age 23 years, SD = 0.86 years) with minimal experience of other cultures (as assessed by questionnaire – see *Appendix A, Screening Questionnaire*). All East Asian validators were Chinese nationals of Chinese heritage, native Chinese speakers who had arrived in the UK for the first time, had a maximum UK residence of 3 months at the time of testing, and possessed a minimum International English Testing System (IELTS) score of 6.0 (Competent user). All Western validators were native English speakers. All validators had normal or corrected-to-normal vision and were free from any lexical, reading, language (e.g., dyslexia) or emotion related atypicalities (e.g., Autism Spectrum Disorder, depression, anxiety) as per self-report, and typically from a student population. I paid each validator £6 per hour, and obtained their written informed consent. The University of Glasgow College of Science and Engineering Ethics Committee authorized the experimental protocol (Ref No: 300150031).

3.4.1.2 Stimuli

For each Western validator, I created a set of stimuli by randomly selecting 245 different models with replacement from the whole set of 490 models derived from the same culture of participants (i.e., Western participants in *Experiment 2*). For each East Asian validator, I created a set of stimuli by randomly selecting 250 different models with replacement from the whole set of 499 models derived from the same culture of participants (i.e., East Asian participants in *Experiment 2*). I displayed each facial expression pattern on one face identity randomly selected with replacement from of 40 new same-race face identities (white Caucasian: 39 European, 1 North American, 20 male, mean age 23 years, *SD* = 5.80 years; East Asian: 40 Chinese, 20 male, mean age 23 years, *SD* = 1.91 years). I then repeated the same procedure for the stimuli used in the incorrect-match trials (see also *Procedure* below) by randomly displaying each facial expression pattern to another face identity. As a result, I produced a total of 490 stimuli for each Western validator (245 models × correct/incorrect matches) and 500 stimuli (250 models × correct/incorrect matches) for each East Asian validator.

3.4.1.3 Procedure

On each experimental trial, I presented a social message label on-screen for 1s, followed by a facial animation. The social message label either correctly or incorrectly matches the following facial expression pattern. After the facial animation, validators responded 'yes' using a keyboard press if the word accurately described the facial animation that followed, and 'no' if it did not. Validators have unlimited time to respond. I assigned yes/no keys to separate hands and counterbalanced key assignments across validators. Half of the trials comprised correct word facial animation matches, with all incorrect matches distributed equally across the other social message categories. Each Western validator therefore completed 490 trials (245 models × correct/incorrect matches) and each East Asian validator completed 500 trials (250 models × correct/incorrect matches), with all trials presented randomly across the experiment for each validator.

I presented all stimuli on a black background displayed on a 19-inch flat panel Dell monitor with a refresh rate of 60 Hz and resolution of 1024 × 1280. Each facial animation (Western: mean size 19.26 cm [SD 1.09] × 12.58 cm [SD 0.38]; Chinese: mean size 17.44 cm [SD 0.96] × 11.45 cm [SD 0.39];) appeared in the validator's central visual field and played only once for a duration of 1.25s. A chin rest ensured a constant viewing distance of 72 cm, with images subtending 15.23° (vertical) and 9.98° (horizontal) of visual angle, reflecting the average size of a human face (Ibrahimagić-Šeper et al., 2006) during natural social interaction (Hall, 1966). I used Matlab 2012a to present stimuli and record responses using a Graphic User Interface.

3.4.1.4 Results

To examine the fundamental components of facial expressions of mental states, I applied a multivariate data reduction analysis called Non-negative Matrix Factorization (NMF, Lee & Seung, 1999) on the facial expression models for each culture separately. NMF ensures that all components and coefficients are non-negative and therefore can accurately reflect the nature of facial movement patterns – each of AU cannot be 'negatively' active. Unlike other factorization analysis such as Principal Component Analysis (PCA, Jolliffe, 2002) or Independent Component Analysis (ICA, Hyvärinen, Karhunen, & Oja, 2004), NMF does not require the main components to be orthogonal or independent and therefore is more appropriate for analyzing facial movement patterns (e.g., Delis et al., 2016). Thus, each facial movement pattern can be represented by how much it relies on the main components (i.e., linear coefficients) and the accents (i.e., the residual) as in (1).

$$\textit{Facial movement pattern} = \textit{Linear coefficients} \times \textit{Main Action Unit patterns} + \textit{Accent} \quad (1)$$

Specifically, for each mental state term I used the facial expression model with the highest d-prime value and those are above perceptual threshold (i.e. d prime > 1). I determined four main components according to the previous study using the same methods on the facial expressions of emotions (Jack et al., 2016). The variance accounted based on these four components is 54%, which is acceptable by the literature (e.g., Peterson, 2000). Figure 3-3 shows the results. For each culture, colour-coded face maps show the main components of facial movement patterns. For each component, color-coded circles on its axis represent the facial expressions of those mental states that rely on each of main component, where the size of the circle and its location on the axis show the weights (i.e., linear coefficients). For each mental state, I plotted its valence according to positive (i.e., magenta) or negative (i.e., cyan). The saturation of the circle represents its level of arousal, where high saturation illustrates high level of arousal. The NMF analysis reveals four main components of facial expressions of mental states and they are common across cultures.

To further demonstrate which social messages are communicated by each of the main components, I examined the main component where each mental state relies on most (i.e., the maximum of linear coefficients) and plotted it according to its level of valence and arousal in the bottom panel of Figure 3-3. The main components are coded using the same color-coding as in the top panel. Specifically, the main component with smiling (AU6, outlined in red) is associated with most of positive mental states. The two main components with Nose Wrinkling (AU9, outlined in blue), Brow Lowerer and Lip Stretcher (AU4 and AU20, outlined in green) are associated with negative mental states. Finally, the component with Brow Raiser, Jaw Drop and Mouth Stretch (AU1-2, AU26 and AU27, outlined in yellow) is primarily associated with high arousal mental states, regardless of its valence.

Interestingly, whereas emotional facial expressions such as ‘thrilled,’ ‘enraged,’ ‘angry,’ ‘disgusted’ and ‘frightened’ heavily rely on one of the main components, conversational facial expressions such as ‘contented,’ ‘encouraged,’ ‘frustrated,’ ‘unsure’ and ‘confused’ seem involve much more complex facial movements, which could either due to the combination of main components or accents with the facial movements that were not characterized by the main components.

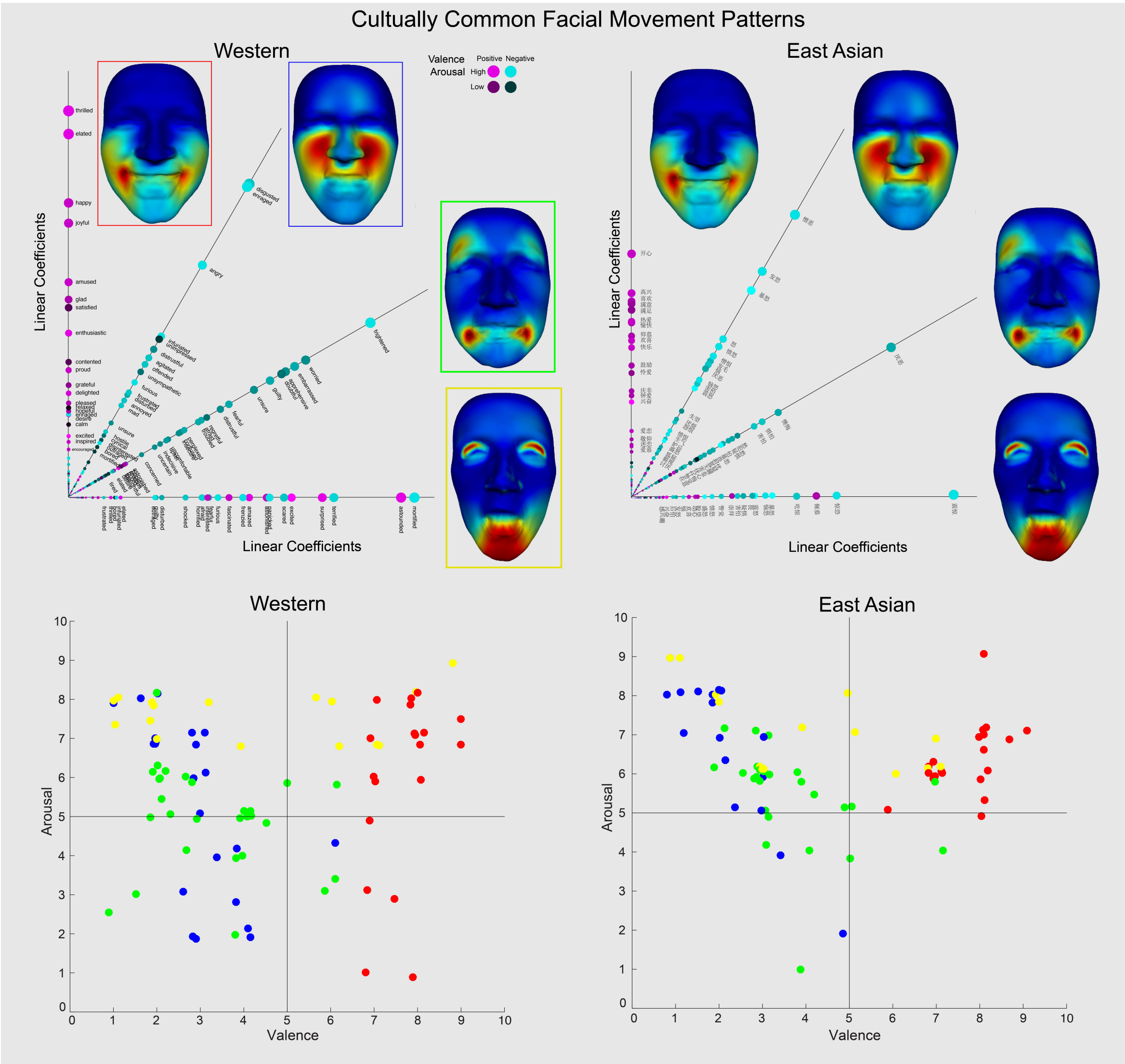

Figure 3-3. Culturally common facial movement patterns of mental states. To extract fundamental facial movement patterns that are used for communicating mental states, I applied Non-negative Matrix Factorization analysis on the facial expression models for each culture separately. Specifically, for each mental state term I used the facial expression model with highest d-prime values and those are above perceptual threshold (i.e. all d prime > 1). For each culture, colour-coded face maps show the main components of facial movement patterns. Results show that four main components are common across cultures. For each component, color-coded circles represent the facial expressions of those mental states that rely on each of main component, where the size of the circle and its location on the axis show the linear coefficients. For each word, I plotted its valence using either magenta (i.e., positive) or cyan (i.e., negative). I plotted the level of arousal of each mental state word using the saturation of the circle, where high saturation represents highly arousal mental states. In the scatter plots on the bottom, each circle shows the level of valence and arousal of each mental state that relies on the main component. The main components are coded using the same color-coding as in the top panel.

Based on the results of the NMF analysis, facial expressions of mental sates can converge across cultures by sharing the main components and accents. In contrast, facial expressions of mental states can diverge across cultures with having distinct main components and/or culture-specific accents. To further explore the cultural convergence and divergence, I examined the 35 pairs of core mental states that have the same meaning in Western and East Asian culture. For each pair of mental states, I examined their main component and their accents across cultures. Figure 3-4 illustrates each condition of cultural convergence and divergence using representative examples. Specifically, facial expressions of ‘troubled/发愁,’ ‘delighted/愉快,’ ‘hostile/憎恨’ and ‘offended/恼’ are culturally common by sharing the main components. Facial expressions of ‘contented/满足’ and ‘enthusiastic/热爱’ are culturally common by sharing the main components and the accents – for example, Cheek Raiser (AU6) is a culturally common accent in ‘contented/满足.’ In contrast, facial expressions of ‘excited/兴奋’ and ‘scared/惧怕’ diverge across cultures. Whereas Westerners use the facial movement patterns of eye whites and widely opening mouth to show exited or scared, East Asian use the facial movement patterns such as smiling or lip stretching that seem less arousal. In addition, facial expressions of mental states with the same main components can also diverge across cultures due to their culture specific accents. For example, facial expressions of ‘sympathetic/怜悯’ rely on the same main component of Brow Lower (AU4) and Lip Stretcher (AU20) across cultures. But Westerns use Dimpler (AU14) and East Asian use Chin Raiser (AU17). Facial expressions of ‘doubtful/怀疑’ have the same main component across cultures. But Westerns tend to use facial movements around the mouth (i.e., horizontally stretching lips), whereas East Asians tend to use facial movements around the eyes (i.e., lowering eye brows).

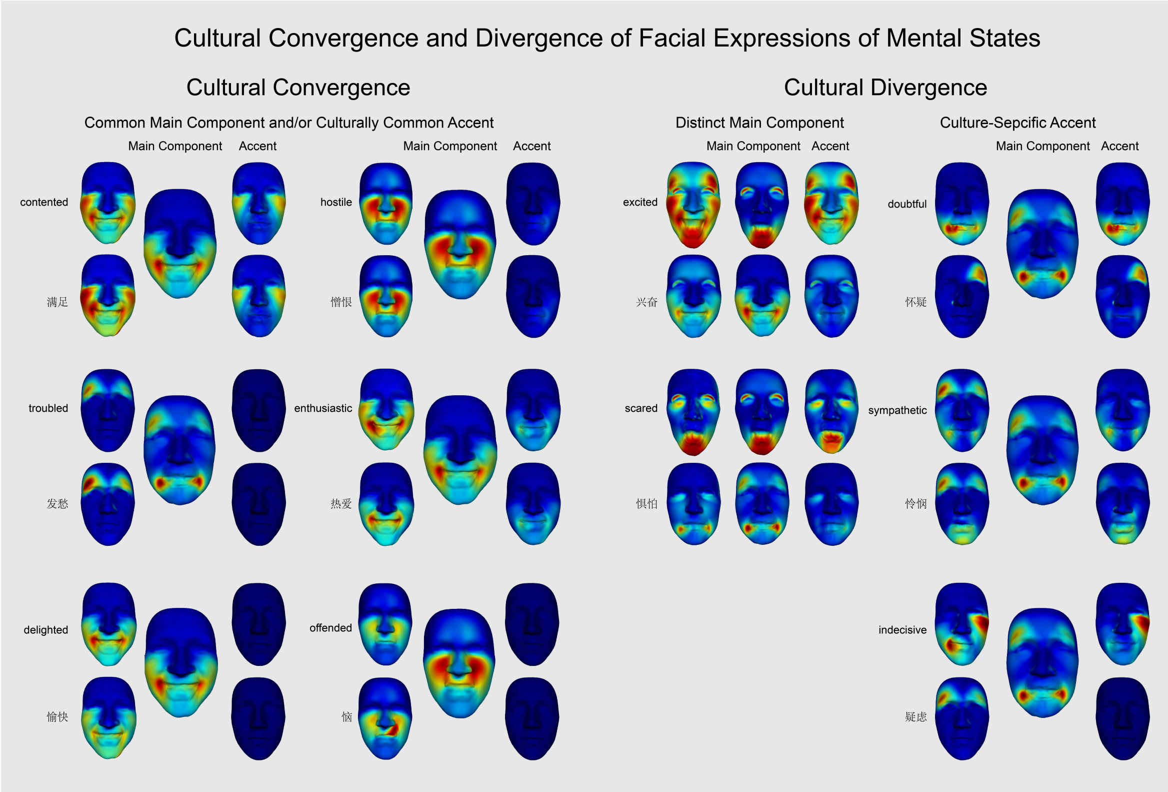

Figure 3-4. Cultural convergence and divergence of facial expressions of mental states. *Culture Convergence*: each set of face maps shows the culturally common main component and accents for each pair of mental states (see labels to left). *Culture Divergence*: each set of face maps shows the distinct main components and/or cultural specific accents (see labels on the top) of each pair of mental states.

3.4.1.5 Discussion

In Study 2, I expanded my research to the facial expressions of a broad set of mental sates in two cultures. Analysis of the resulting models reveals four culturally common facial movement patterns, which supports the hypothesized fundamental facial movement patterns by the previous study on the facial expressions of emotions (Jack et al., 2016). However, whereas emotional facial expressions primarily rely on one of main component, conversational facial expressions have more complex facial movement patterns. Cross-cultural comparison reveals the source of cultural convergence/divergence of facial expressions. Specifically, facial expressions can converge across cultures by sharing main components and accents. In contrast, facial expressions can diverge cross cultures by either having completely different components or culture-specific accents.

Culturally common facial movement patterns for communicating valence and arousal

The NMF analysis demonstrated that four main facial movement patterns are common across culture, which support the theory that these four facial movement patterns are fundamental facial expressions that are used for communicating social messages across cultures (Jack et al., 2016). Interestingly, my results show that facial expressions of positive mental states are much similar to each other and therefore primarily reply on one component (i.e., smiling), whereas negative mental states are communicated using more distinct facial movement patterns. This corresponds the results in Experiment 1, which shows that the number of negative mental sates is much more than the positive mental states in both cultures. Previous studies on threat-related facial expressions show that observers responded faster and more accurately on detecting such facial expressions based on both behaviour data (e.g., Öhman, Lundqvist, & Esteves, 2001) and neural activity (Anderson, Christoff, Panitz, De Rosa, & Gabrieli, 2003; Morris, Frith, Perrett, & Rowland, 1996). Thus, more distinctive facial movement patterns may have evolved to communicate negative mental states due to the importance of discriminating such negative social messages – for example, between threat and non-threat messages. In more general context, this reflects the stronger impact of 'bad' events/states on human daily life and the long-term adaption from avoiding harmful consequences in human evolution (see a review of Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001).

Finally, both culture use opening eyes and mouth for communicating highly arousal mental states. Whereas Westerns use this facial movement pattern to communicate both positive mental sates (e.g., 'excited,' 'amazed') and negative mental states (e.g., 'terrified,' 'scared'), East Asians tend to associate the highly arousal facial expressions with negative social messages. For example, similarly as the facial expressions of positive physical states (i.e., 'orgasm') in Study 1, East Asians use much calmer smiling to show 'excited.' This supports the previous studies suggesting that Westerns prefer experiencing and expressing high arousal states, whereas East Asians prefer much calmer states (e.g., Lim, 2016; Tsai et al., 2006).

Cultural divergence of conversational facial expressions

Although both emotional facial expressions and conversational facial expressions show convergence and divergence across cultures, conversational facial expressions tend to less rely on main components and have more complex facial movements – for example, accents that have facial movements are not characterized by the main components. Specifically, the four components may have evolved based on their physiological origins as sensory regulator (e.g., Chapman et al., 2009; Lee et al., 2014; Susskind et al., 2008) and therefore those emotional facial expressions that primarily rely on such components are more likely to converge across cultures. In contrast, conversational facial expressions have more accents that could have evolved for the function of social communication (Shariff & Tracy, 2011) and therefore diverge across cultures (Darwin, 1872/1998).

To further explore the role of culture in impacting social communication, I conducted a cross-cultural study using the four key conversational facial expressions – 'interested,' 'bored,' 'confused' and 'thinking' in Study 3.

4 Study 3: Exploring The Dual Role of Culture on Signalling and Receiving Four Key Conversational Facial Expressions

4.1 Introduction

Humans are a highly social species and frequently engage in complex interactions with a variety of people to negotiate the different facets of life (Wilson, 2012). One of the most effective ways that humans can communicate is by using facial expressions (see also voice, hand gesture, and body posture, e.g., Aviezer et al., 2012; Kendon, 2004; Krahmer, Ruttkay, Swerts, & Wesselink, 2002; Schegloff, 1982). Within a society, knowledge of which facial expressions represent which social messages is shared across its members, thus enabling clear communication (but see also Adolphs, Tranel, Damasio, & Damasio, 1994; Celani et al., 1999; Kohler et al., 2003; Shannon, 1948/2001). In contrast, communication breakdown is most likely to occur across groups where such expectations differ, such as facial expressions of emotions across distinct cultures (e.g., Chen et al., 2016; Elfenbein et al., 2007; Jack, Blais, Scheepers, Schyns, & Caldara, 2009; Jack, Caldara, et al., 2012; Kinloch & Metge, 2014; LaBarre, 1947) (but see also Darwin, 1872/1998 for cultural commonalities; Ekman, 1971; Ekman et al., 1987; Ekman et al., 1969; Izard, 1971). Given the critical importance of social communication for successful societal functioning, understanding the sources of communication success and failure within and between cultures remains a central goal in psychology, with recent impact in computer vision and social robotics. Here, I examine a pervasive component of communication that plays a substantial role in comprehension during most face-to-face social exchanges – conversational facial expressions – and show specifically how culture impacts two sources of communication clarity – signalling structure and receiver response.

Conversational facial expressions such as of interest and confusion play a critical role in communication clarity during social exchanges by guiding the flow and content of the interaction (e.g., Bavelas & Chovil, 1997; Chovil, 1991; Eibl-Eibesfeldt, 1989; Izard, 2013). For example, certain conversational facial expressions can alter the meaning of spoken words (e.g., Bavelas & Chovil, 2000), indicate when and who will take turn (e.g., Bull & Frederikson, 2001; Duncan, 1972; Yngve, 1970), or reflect the impact of the sender's message on others including interest, boredom, or confusion (e.g., Cunningham, Kleiner, Bülthoff, & Wallraven, 2004; El Kaliouby & Robinson, 2005). Yet, little is known about whether and how the specific structure of these signals induces clear understanding within and across cultures or causes confusions (e.g., see Cunningham & Wallraven, 2009; Darwin, 1872/1998; Ekman, 1979; Ekman & Friesen, 1977; Hass, 1970; Izard, 2013; Nusseck et al., 2008). For example, although Westerners use specific eyebrow and mouth movements to communicate conversational messages such as interest and confusion (e.g., Cunningham et al., 2004), it remains unknown whether these specific signalling structures with hypothesized physiological origins (e.g., Darwin, 1872/1998; Ekman, 1979) generalize across cultures. Signal evolution accounts predict that facial expressions used for social communication could comprise both culturally common iconic face movements, such as wrinkling the nose to show rejection (Bavelas & Chovil, 1997; Chapman et al., 2009; Rozin, Haidt, & McCauley, 2008), and culture-specific accents that might be learned and therefore hinder cross-cultural communication (Chen et al., 2016; Elfenbein, 2013; Elfenbein et al., 2007; e.g., Marsh, Elfenbein, & Ambady, 2003; see Slabbekoorn & den Boer-Visser, 2006 for accents in birdsong). Consequently, a precise knowledge of signal structures within and across diverse cultures is critical to understanding the sources of human communication success and failure with direct impact in related fields such as social robotics.

I address this first issue by objectively characterizing the specific dynamic facial expression patterns that communicate four key conversational messages – 'interested,' 'bored,' 'confused,' and 'thinking' – in Western and East Asian culture (Preliminary Study) using 4D computer graphics and a data-driven psychophysical approach. I specifically chose these four conversational messages because they perform the pervasive social function of guiding the flow of interactions such as calling for more information, re-routing, pausing, or terminating an interaction (e.g., Bavelas et al., 2000; Cassell et al., 2000; Chovil, 1991; Cunningham & Wallraven, 2009; Ekman, 2004; El Kaliouby & Robinson, 2005; Kaulard, Cunningham, Bülthoff, & Wallraven, 2012; Wallraven et al., 2015). To understand the role of signal structure on communication, I first conducted a cross-cultural communication task with a new set of Western and East Asian observers (Experiment) to measure communication clarity. To analyze and compare the facial expression signalling structures of each culture, I then used a new information theoretic approach and isolated the face movement patterns that are common across cultures and could facilitate communication, from those that are culture specific and could hinder communication. Next, to understand the second role of culture, I analysed the response characteristics of observers in each culture – i.e., the impact of the information pattern on the receiver's behaviour (e.g., Dukas, 1998; Scott - Phillips, 2008; Shannon, 1948/2001). Here, I identified specific biases in the response of observers that could increase or decrease communication clarity within and across cultures. My results reveal that culture plays a dual role in communication clarity by shaping both the structure of conversational facial expression signals and the response characteristics of the observer.

4.2 Experiment 1: Characterizing Dynamic Conversational Facial Expressions in Western and East Asian Culture

To examine whether conversational facial expressions are understood across cultures, I must first identify which facial expression patterns communicate conversational messages within each culture (here, Western and East Asian). To this aim, I reverse engineered a set of such dynamic conversational facial expressions by proceeding in two steps. First, I used a data-driven, social psychophysics approach based on reverse correlation that combines random face movements with subjective human perception to characterize the dynamic facial expression patterns that communicate specific social messages to individuals in a given culture. I then validated each resulting facial expression pattern within each culture using a second set of observers. Figure 4-1 illustrates my data-driven method using one example trial.

4.2.1 Method: Data-driven, social psychophysics approach

On each experimental trial, a dynamic facial expression generator (Yu et al., 2012) randomly selected a set of individual face movements called Action Units (e.g., AUs, Ekman & Friesen, 1976) from a core set of 42 AUs. For example, in Figure 4-1 the selection comprises Outer Brow Raiser (AU2) color-coded in blue, Upper Lid Raiser (AU5) in green, and Lip Corner Puller (AU12) in red. A random movement is then applied to each AU separately using random values selected for each of six temporal parameters (i.e., onset latency, acceleration, peak amplitude, peak latency, deceleration, and offset latency; in Figure 4-1 see labels illustrating the blue curve). In Figure 1, the color-coded activation curves next to each AU illustrate its randomly assigned movement. These dynamic AUs are then combined to produce a photo-realistic facial animation, shown in Figure 1 with four snapshots across time for an animated example; see also *Stimulus Generation* for further details). The observer viewed the random facial animation and categorized it according to one of the four conversational messages – i.e., 'interested,' 'bored,' 'confused,' or 'thinking' – if the random facial movements formed a pattern that communicated (i.e., elicited the perception of) that message to that observer (e.g., in Figure 4-1, 'interested'). The observer also rated the intensity of the message perceived on a 5-point scale from 'very weak' to 'very strong' (e.g., in Figure 4-1, 'strong'). If the face movements did not communicate any one of the four conversational messages, or conveyed another message (e.g., a blend such as 'interestedly bored,' Du, Tao, & Martinez, 2014) the observer selected 'other' – a response category that avoids the limitations of classic alternative forced-choice tasks (see Russell, 1994 for discussion). Each of 20 Western and 20 East Asian observers (see *Participants* below) categorized 2,400 such facial animations displayed on same-race faces (8 white Caucasian, 4 male, mean age 23 years, *SD* = 4.1 years; 8 Chinese, 4 male, mean age 22 years, *SD* = 1.0 years). I presented the response options in the observer's native language – i.e., English or simplified Chinese (a professional translator provided the Chinese words and confirmed that each perfectly matched the meaning of its English word counterpart; see full details in *Procedure*). Thus, this method can capture the dynamic facial expression patterns that communicate different social messages to individuals in a given culture (see Gill et al., 2014; Jack, Garrod, et al., 2012; Richoz et al., 2015, for further examples).

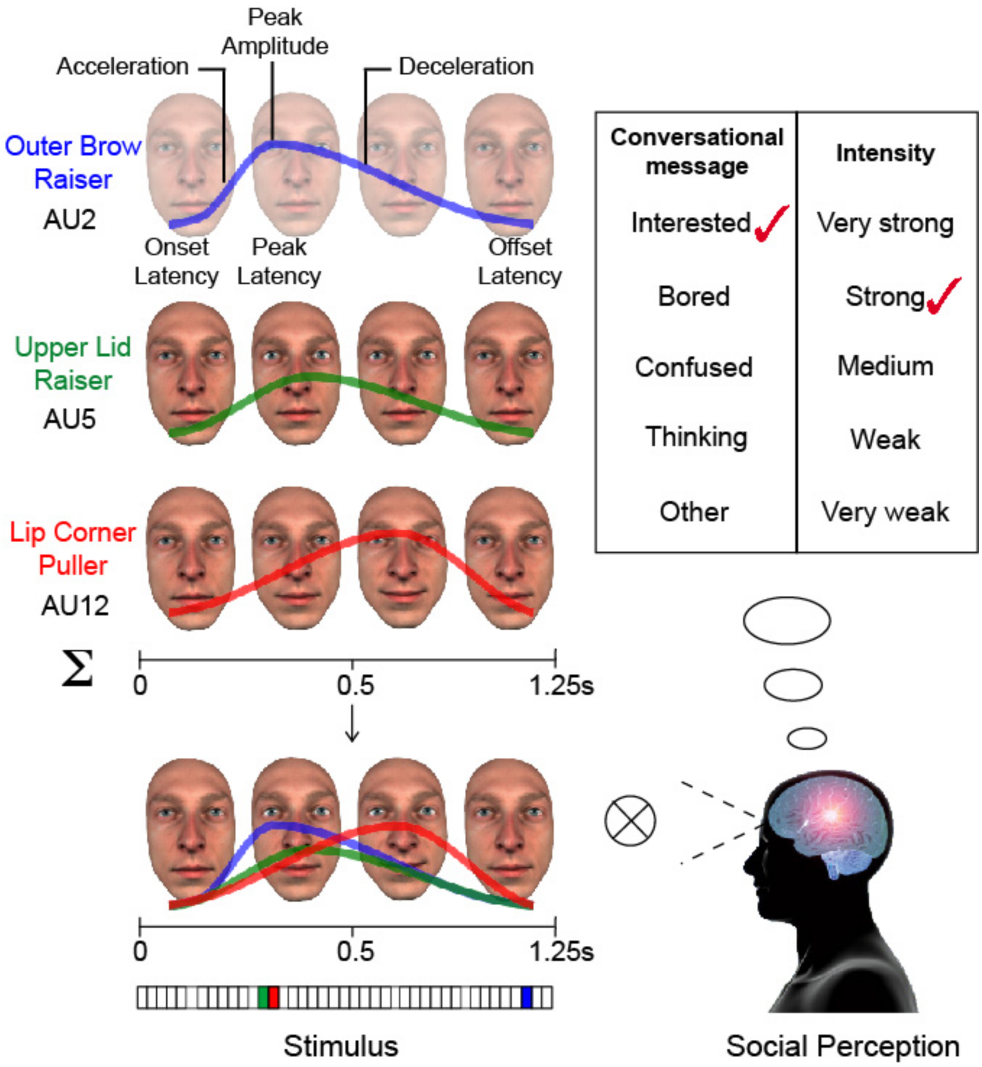

Figure 4-1. Characterizing dynamic facial expressions of conversational messages. *Stimulus.* On each experimental trial, a dynamic facial expression generator (Yu et al., 2012) randomly selected a biologically legitimate combination of individual facial movements called Action Units (e.g., AUs, Ekman & Friesen, 1976) from a core set of 42 AUs (here, Outer Brow Raiser – AU2 color-coded in blue, Upper Lid Raiser – AU5 in green, and Lip Corner Puller – AU12 in red). A random movement is applied to each AU separately using random values selected for each of six temporal parameters (i.e., onset latency, acceleration, peak amplitude, peak latency, deceleration, and offset latency; see labels illustrating the blue curve). The randomly activated AUs are then combined to produce a facial animation, shown here with four snapshots across time. The color-coded vector below shows the three randomly selected AUs on this example trial. *Social Perception.* The observer views the facial animation and categorized it according to one of the four conversational messages – i.e., 'interested,' 'bored,' 'confused,' or 'thinking' – if the random animation formed a pattern that communicated that message to them (e.g., 'interested'). The observer also rated the intensity of the message perceived on a 5-point scale (e.g., 'strong'). Otherwise, the observer selected 'other.' Each of 20 observers in each culture categorized 2,400 such facial animations displayed on same-race faces.

4.2.1.1 Participants

I recruited 20 Western white Caucasian (10 male, mean age 21 years, *SD* = 2.3 years) and 20 East Asian observers (Chinese nationality and heritage, 10 male, mean age 23 years, *SD* = 2.1 years). To control for the possibility that any culturally common facial expression patterns could have been learned from cross-cultural interactions, I recruited observers with minimal exposure to and engagement with other cultures (De Leersnyder et al., 2011) as assessed by screening questionnaire (see *Appendix A, Screening Questionnaire*). All East Asian observers had a maximum UK residence of 3 months at the time of testing, and had a minimum International English Testing System (IELTS) score of 6.0 (Competent User). All observers had normal or corrected-to-normal vision and were free from any lexical, reading, language (e.g., dyslexia) or emotion related atypicalities (e.g., Autism Spectrum Disorder, depression, anxiety) as per self-report. I paid each observer £6 per hour, and obtained their written informed consent prior to testing. The University of Glasgow College of Science and Engineering Ethics Committee authorized the experimental protocol (Ref No: 300140082).

4.2.1.2 Stimulus Generation

On each experimental trial, a dynamic facial expression generator (Yu et al., 2012) – randomly selected a set of individual face movements called Action Units (AUs, Ekman & Friesen, 1976) from a core set of 42 AUs using a binomial distribution (with the two parameters of the binomial distribution set to $n = 5$ and $P = 0.6$). This random selection of AUs on each trial produced a median of 3 AUs across all experimental trials. Figure 4-1 illustrates the stimulus generation procedure using one illustrative trial. A random movement is then assigned to each of the randomly chosen AUs by selecting random values (from a uniform distribution over the interval [0,1]) for each of six temporal parameters (i.e. onset latency, acceleration, peak amplitude, peak latency, deceleration, and offset latency; In Figure 1, see labels illustrating the blue curve). The time courses for each AU are generated using a smooth curve (i.e., a cubic Hermite spline, Yu et al., 2012) as shown by the color-coded activation curves in Figure 4-1. The randomly activated AUs are then combined to produce a photo-realistic facial animation lasting 1.25s, shown in Figure 1 with four snapshots across time. I displayed each facial animation on one of 8 white Caucasian (8 British, 4 male, mean age 23 years, *SD* = 4.1 years) or East Asian (8 Chinese, 4 male, mean age 22 years, *SD* = 1.0 years) face identities, and generated all stimuli using a standard procedure in 3D Studio Max.

4.2.1.3 Procedure

Each observer categorized the same 2,400 random facial animations displayed on same-race faces, presented in random order across the experiment. I presented all stimuli on a black background displayed on a 19-inch flat panel Dell monitor with a refresh rate of 60 Hz and resolution of 1024 × 1280. Each stimulus (mean size: 19 cm [SD 0.83] × 12 cm [SD 0.49]) appeared in the observer’s central visual field and played only once for a duration of 1.25s. A chin rest ensured a constant viewing distance of 71 cm, with images subtending 15.24° (vertical) and 9.66° (horizontal) of visual angle, which reflects the average size of a human face (Ibrahimagić-Šeper et al., 2006) during natural social interaction (Hall, 1966). I recruited a professional translator to provide Chinese words for 'interested' (感兴趣), 'bored' (无聊), 'confused' (困惑), and 'thinking' (思考). To establish that the Chinese words have the same meaning as in English, I asked the professional translator to rate similarity of meaning of each Chinese word provided for each English word on a scale of 1 – 5 where 5 is a perfectly matched meaning. The translator rated each pair of words listed above as 'perfectly matched' (i.e., rating of 5). I used Matlab 2012 to present the stimuli and record responses using a Graphic User Interface.

4.2.1.4 Results

Following the experiment, I used an established reverse correlation analysis to identify the dynamic AUs that are significantly correlated with the perception of each conversational message for each individual observer in each culture.

Reverse Correlation Analysis. To identify the dynamic AUs that are significantly correlated with (i.e., elicit the perception of) each conversational message – 'interested,' 'bored,' 'confused,' and 'thinking'– for each individual observer, I used an established procedure as follows. First, I performed a Pearson correlation for each individual AU between two binary vectors – the first vector recorded the presence vs. absence of the AU considered on each trial; the second vector recorded the responses of the observer on each corresponding trial (e.g., 'interested'). For all AUs significantly correlated with a given response, I assigned a value of 1 (two-tailed $p < .05$) and 0 otherwise, resulting in a 1×42-dimensional binary vector. Thus, each resulting vector details the AUs that are significantly correlated with the perception of each conversational message for each individual observer. I did not analyze trials categorized as 'other' because they do not correspond to any specific social message.

For each significant AU in the 1×42-dimensional binary vector, I also computed an estimate of its temporal dynamics as follows. For each of the six temporal parameters (i.e., onset latency, acceleration, peak amplitude, peak latency, deceleration, and offset latency, cf. labels illustrating the blue curve in Figure 1), I performed an independent linear regression between the observer's intensity ratings (e.g., 'very strong') and all trials where the observer selected the conversational message in question (e.g., 'interested'). Thus, I computed a total of 160 dynamic facial expression patterns (20 observers × 4 conversational messages × 2 cultures) each represented as a 1 × 42-dimensional binary vector detailing the significant AUs, plus 6 values detailing the temporal parameters of each significant AU. Computing these dynamic facial expression patterns in this way enables their reconstruction as stimuli for subsequent behavioral tasks. To derive movies of the resulting dynamic facial expression patterns to use as stimuli, and for illustration purposes, I combined the significantly correlated AUs with their corresponding temporal parameters derived from the regression coefficients and displayed the dynamic pattern on different face identities. For stimuli, I used the facial expression patterns derived from 'high intensity' ratings as these represent the most salient signals.

I thus produced a total of 160 statistically robust dynamic facial expression patterns (20 observers × 4 conversational messages × 2 cultures) each represented as a 1 × 42 dimensional vector detailing the composition of significant AUs along with 6 values detailing the movement pattern of each significant AU. Figure 4-2 shows the distribution of significant AUs across all individual observer facial expression patterns in each culture.

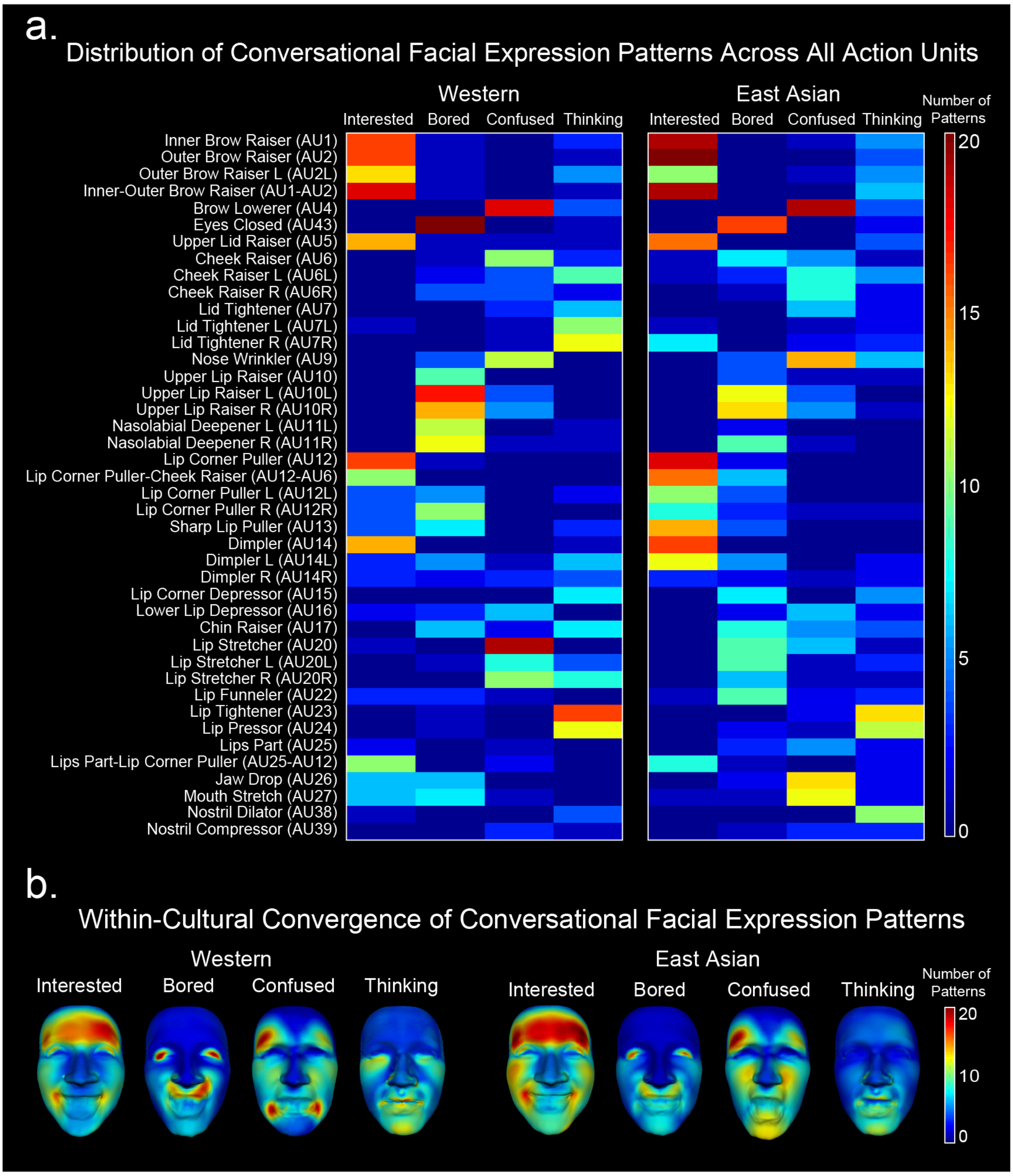

Figure 4-2. Distribution of conversational facial expression patterns across all Action Units and their convergence within Western and East Asian culture. (a) *Distribution of conversational facial expression patterns across all Action Units.* Color-coded matrices show for each culture – Western (left panel) and East Asian (right panel) – and the four conversational messages – 'interested,' 'bored,' 'confused,' and 'thinking' – the number of individual observer facial expression patterns (maximum = 20) with each Action Unit (AU; see labels on left) as significant. Warmer colors indicate a higher number of patterns; cooler colors indicate a lower number (see colorbar on right). For example, all Western facial expressions of 'confused' include Lip Stretcher (AU20) as indicated by the red coloring but none included Jaw Drop (AU26) as indicated by the dark blue coloring. (b) *Within-cultural convergence of conversational facial expression patterns.* Color-coded face maps show for each culture and conversational message separately the number of individual observer facial expression patterns with each AU using the same color-coding as in panel (a).

4.3 Experiment 2: Within-Culture Validation of Dynamic Conversational Facial Expressions

Next, I validated each of the dynamic facial expression patterns by measuring its within-culture communication accuracy using d-prime (Green & Swets, 1966). D-prime provides a reliable measure of communication accuracy as it considers both hit rates (signal present, 'yes' response) and false alarm rates (signal absent, 'yes' response), which safeguards against artificially high accuracy rates (e.g., see Elfenbein et al., 2002; Lynn & Barrett, 2014; Russell, 1994) D-prime also provides a measure that reflects key and independent components of communication – signal discriminability (e.g., see Guilford & Dawkins, 1991) and biases in the response characteristics of the observer (e.g., see Stanislaw & Todorov, 1999).

4.3.1 Methods

4.3.1.1 Participants

To validate the facial expressions within each culture, I recruited a new set of 18 Western white Caucasian (8 male, mean age 21 years, *SD* = 2.9 years) and 12 East Asian validators (Chinese nationality and heritage, 2 male, mean age 23 years, *SD* = 0.8 years) using the same criteria as detailed above.

4.3.1.2 Procedure

Each validator completed a signal detection task using the facial expression patterns derived from their own culture and displayed on one of 4 new same-race faces (white Caucasian: 2 male, mean age 25 years, *SD* = 4.4 years; East Asian: 2 male, mean age 23 years, *SD* = 3.5 years). On each experimental trial, validators viewed one of the four conversational messages – 'interested,' 'bored,' 'confused,' or 'thinking' – on-screen for 2 seconds followed by either a correctly or incorrectly matched dynamic facial expression pattern lasting 1.25s. Validators responded 'yes' using a keyboard press if the word accurately described the facial animation or 'no' if it did not. Half of the trials comprised correct label-facial expression matches and included all 320 facial expressions with all incorrect matches distributed equally across the other conversational messages. Each validator therefore completed 640 trials (20 facial expressions × 4 conversational messages × 4 same-race identities × correct/incorrect matches) with all trials presented randomly across the experiment. I assigned yes/no keys to separate hands and counterbalanced key assignments across validators. Thus, I presented each facial expression pattern more than 75 times across the experiment in each culture, which is sufficient to obtain a statistically robust measure of signal sensitivity – i.e., d-prime (Stanislaw & Todorov, 1999).

4.3.2 Results

For each individual facial expression pattern I computed d-prime (Green & Swets, 1966) by pooling the responses of all validators in each culture separately. Figure 4-3 below shows all individual facial expression patterns displayed as face maps and ranked by d-prime.

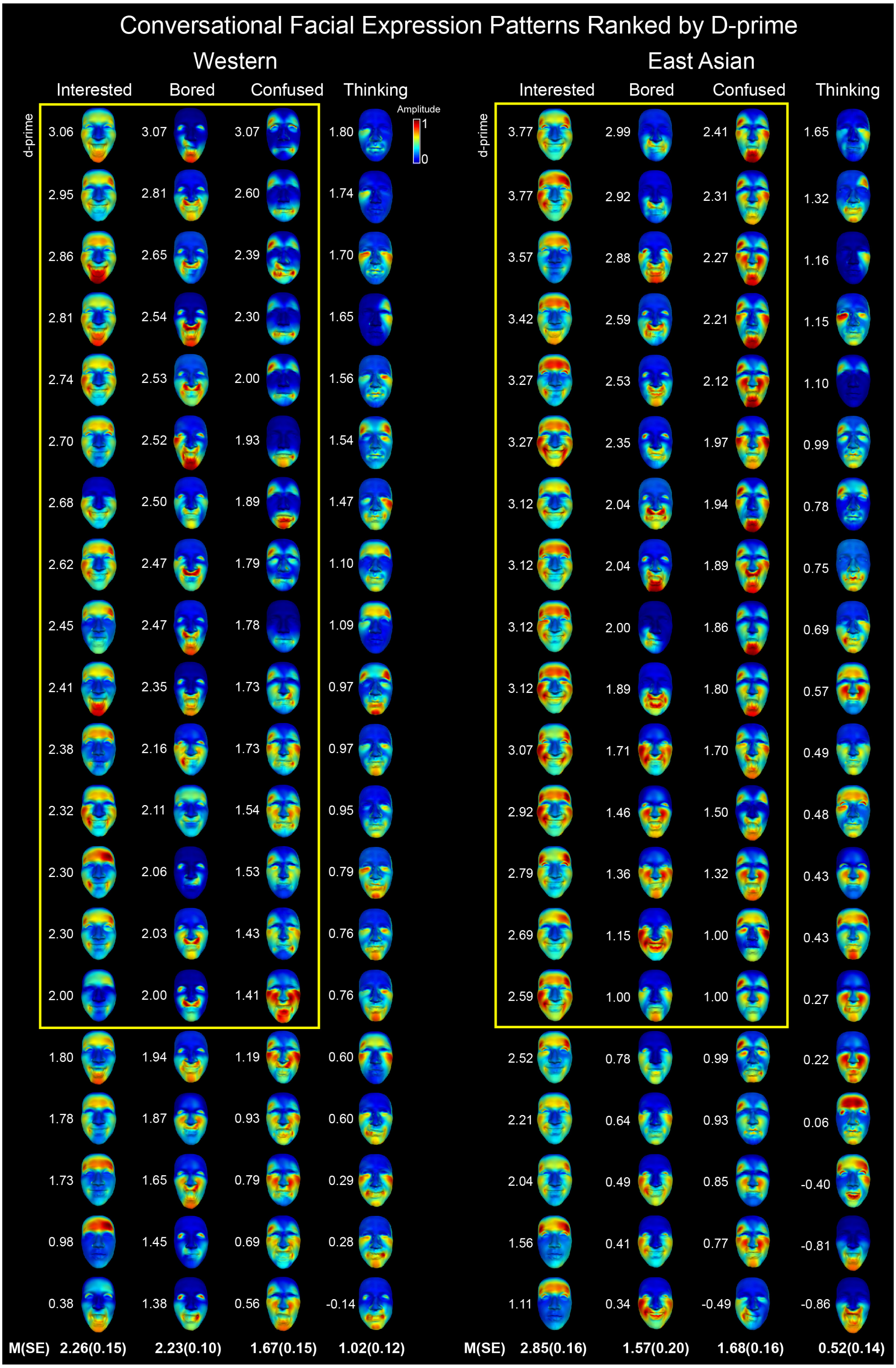

Figure 4-3. Conversational facial expression patterns ranked by d-prime in each culture. Color-coded face maps show for each culture – Western (left panel) and East Asian (right panel) – the individual observer facial expression patterns of each conversational message. Color-coding represents the amplitude (i.e., intensity) of each significant Action Unit (see colorbar in center) as derived from the reverse correlation analysis. Facial expression patterns are ranked according to d-prime (shown to the left of each face map) in

descending order from top to bottom. Below each set of conversational face maps is the corresponding average d-prime values and standard error of the mean. Facial expression patterns outlined in yellow show those with the highest 15 d-prime values for each conversational message, which I retained for use as stimuli in the cross-cultural communication task and further analysis.

For further cross-cultural testing, and to ensure equal numbers of facial expressions per conversational message, I retained the top 75% of facial expression patterns of each conversational message (N = 15) in each culture if they all exceeded perceptual threshold (i.e., d-prime > 1). Since fewer than 75% of 'thinking' facial expressions exceeded perceptual threshold in each culture (Western: 45%, N = 9, East Asian: 25%, N = 5), I retained the top 75% of facial expressions of 'interested,' of 'bored,' and of 'confused' (d-prime ranged from 1.00 to 3.77). Figure 4-4 shows the within-cultural variance of these facial expression patterns.

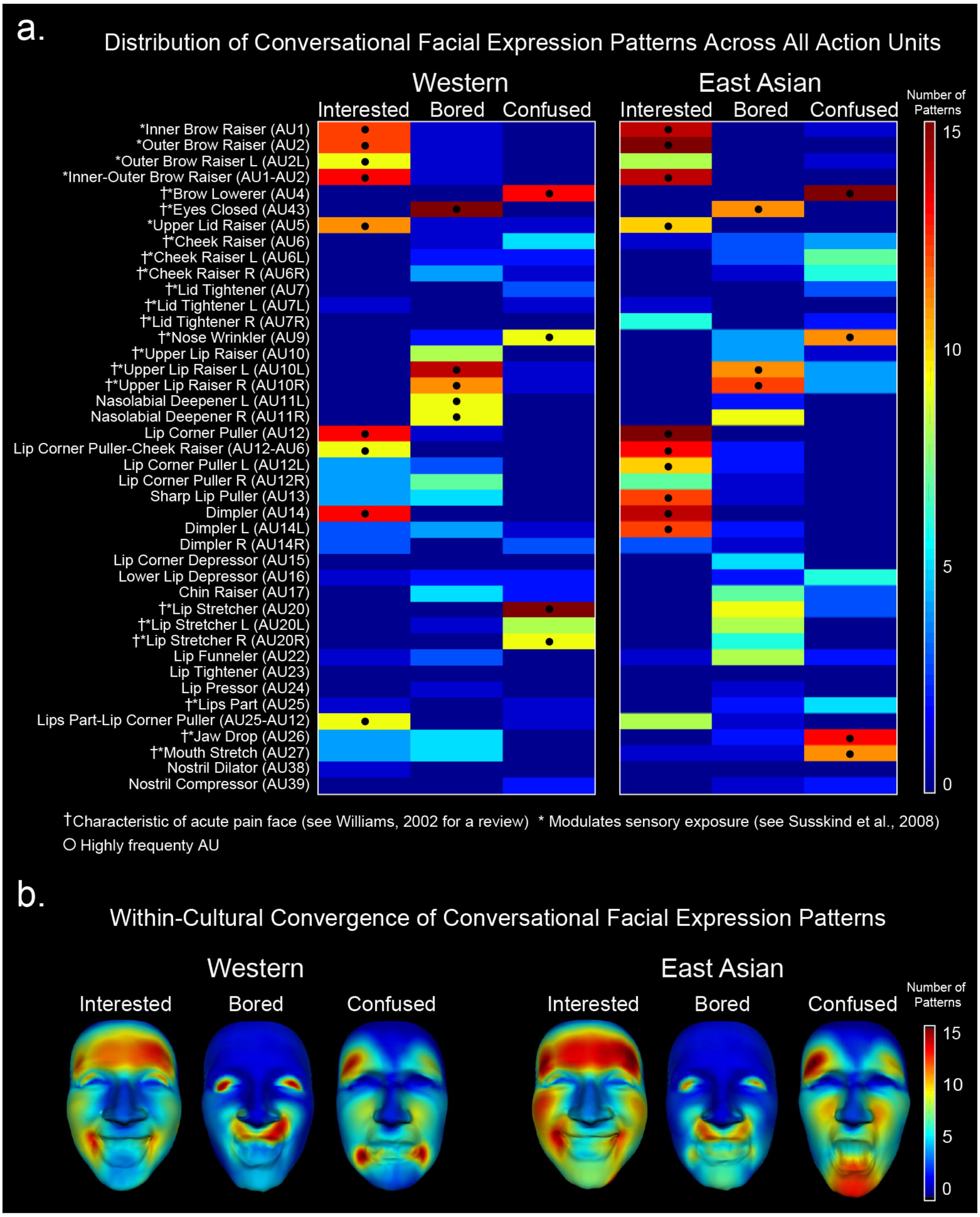

Figure 4-4. Within-cultural convergence of conversational facial expression patterns used in the cross-cultural communication task. (a) *Distribution of conversational facial expression patterns across all Action Units.* Color-coded matrices show for each culture – Western (left panel) and East Asian (right panel) – and the three conversational messages – 'interested,' 'bored,' and 'confused' – the number of individual observer facial expression patterns (maximum = 15) with each Action Unit (AU; see labels on left) as significant. Warmer colors indicate a higher number of patterns; cooler colors indicate a lower number (see colorbar on right). For example, all Western facial expressions of 'bored' include Eyes Closed (AU43) as indicated by the dark red coloring but is not present in any 'interested' or 'confused' facial expressions as indicated by the dark blue coloring. AUs marked with † indicate those characteristic of acute pain facial expressions (Williams, 2002); AUs marked with an asterisk (*) indicate those that modulate sensory exposure (Susskind et al., 2008); AUs marked with a black dot are *highly frequent* across the facial expressions patterns. (b) *Within-cultural convergence of conversational facial expression patterns.* Color-coded face maps show for each culture and conversational message separately the number of individual observer facial expression patterns with each AU as significant using the same color-coding as in panel (a).

4.4 Experiment 3: Cross-Cultural Communication Clarity of Conversational Facial Expressions

My primary aim is to understand whether conversational facial expressions are understood across cultures or if they produce cross-cultural confusions. To this aim, I displayed the set of culturally valid dynamic facial expressions derived in the preliminary study above on a new set of face identities and conducted a cross-cultural communication task with a new set of Western and East Asian observers.

4.4.1 Methods

4.4.1.1 Participants

I recruited a new group of 20 Western white Caucasian (10 male, mean age 21 years, *SD* = 3.3 years) and 20 East Asian (Chinese nationality and heritage, 10 male, mean age 24 years, *SD* = 1.8 years) observers using the same criteria as detailed above.

4.4.1.2 Stimuli

For each observer, I displayed each culturally validated facial expression pattern on 4 new same- and other-race face identities (2 white Caucasian, 1 male; 2 East Asian, 1 male) pseudo-randomly selected (with replacement) from a set of 8 white Caucasian identities (4 male, mean age 23 years, *SD* = 2.4 years) and 8 East Asian identities (8 Chinese, 4 male, mean age 21 years, *SD* = 1.4 years). Thus, for each observer I generated a set of 360 facial animations (15 facial expression patterns × 3 conversational messages × 2 cultures × 4 same- and other-race face identities).

4.4.1.3 Procedure

On each experimental trial, observers viewed one of the three conversational messages – ‘interested,’ ‘bored,’ or ‘confused,’ – on-screen for 2 seconds followed by either a correctly or incorrectly matched facial expression. Observers responded ‘yes’ using a keyboard press if the label accurately described the facial animation that followed, and ‘no’ if it did not. I assigned yes/no keys to separate hands and counterbalanced key assignments across observers. Half of the trials comprised correct message-facial animation matches and included all 360 facial animations, with all incorrect matches distributed equally across the other conversational messages. Each observer therefore completed 720 trials (360 stimuli × 2 correct/incorrect matches) comprising same- and other-culture facial expression patterns. I thus presented each facial expression pattern more than 75 times across the experiment in each culture, which is sufficient to obtain a statistically robust measure of signal sensitivity – i.e., d-prime (Stanislaw & Todorov, 1999).

4.4.2 Results

4.4.2.1 Cross-Cultural Analysis of Communication Clarity

Following the experiment, I computed d-prime for each facial expression pattern within and across cultures separately by pooling the responses of all observers in each culture. Figure 4-5a shows the results for each conversational message and culture. In each plot, color-coded points represent individual facial expression patterns derived from a cultural observer in the preliminary study (magenta – Western; green – East Asian; see legend on left) plotted according to its communication clarity (i.e., d-prime) within and across cultures (see observer labels on x axis). Color-coded lines represent the average differences in the communication clarity of the facial expression patterns across cultures (see Table 4-1 for corresponding average d-prime values).

Table 4-1

Communication Clarity (d-prime) of Conversational Facial Expression Patterns Within and Across Cultures

	Within-Culture		Cross-Culture	
Culture of Observer	Western	East Asian	Western	East Asian
Culture of Facial Expression	Western	East Asian	East Asian	Western
Facial Expression	*Mean (SE)*	*Mean (SE)*	*Mean (SE)*	*Mean (SE)*
Interested	2.42 (0.12)	2.28 (0.09)	2.47 (0.08)	1.88 (0.14)
Bored	1.86 (0.09)	1.40 (0.10)	1.61 (0.13)	1.21 (0.09)
Confused	2.22 (0.08)	1.44 (0.07)**	1.17 (0.13)**	0.82 (0.11)**

Note. ** $p < .001$

I then applied to these d-prime values a 2 (culture of facial expression pattern) × 2 (culture of observer) analysis of variance (ANOVA). A power analysis[1] using G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007) based on the estimated effect size (i.e., η_p^2 = 0.14) for our analysis (i.e., ANOVA, see Cohen, 1988; Cohen, 1992 for more details) and a significance level of 0.05 confirmed that the sample size used (N = 15) is sufficient for high power (i.e., estimated power is 0.82). Analysis revealed that facial expressions of 'interested' and 'bored' elicited similarly high communication clarity within and across cultures (i.e., no significant interaction in either culture; ANOVA, F (1, 56) = 2.60, two-tailed p = .11, η_p^2 = .04 for 'interested;' F (1, 56) = 3.21, two-tailed p = .08, η_p^2 = .05 for 'bored'). Western observers showed higher performance compared to East Asian observers for facial expressions of 'interested' (F (1, 56) = 11.44, two-tailed p = .001) and 'bored' (F (1, 56) = 16.18, two-tailed p < .001) but with relatively small effect sizes (η_p^2 = 0.16, η_p^2 = 0.20, respectively). In contrast, facial expressions of 'confused' showed a significant within-culture advantage in both cultures (ANOVA, F (1, 56) = 71.52, two-tailed p < .001, η_p^2 = .56; Western: two-tailed p < .001, Cohen's d = 1.58, 95% CI [0.66, 1.43]; East Asian: two-tailed p < .001, Cohen's d = 1.34, 95% CI [0.20, 1.00]) with significantly higher within-culture performance for Western than East Asian facial expressions (pairwise comparison with Bonferroni correction, two-tailed p < .001, Cohen's d = 1.56, 95% CI [0.40, 1.16]; no significant difference for cross-cultural performance; two-tailed p = .08, Cohen's d = .70, 95% CI [-0.03, 0.74]).

[1] It should be noted that since each observation in the ANOVA analysis is the responses to one facial expression model, the sample size of power analysis is the number of facial expression models.

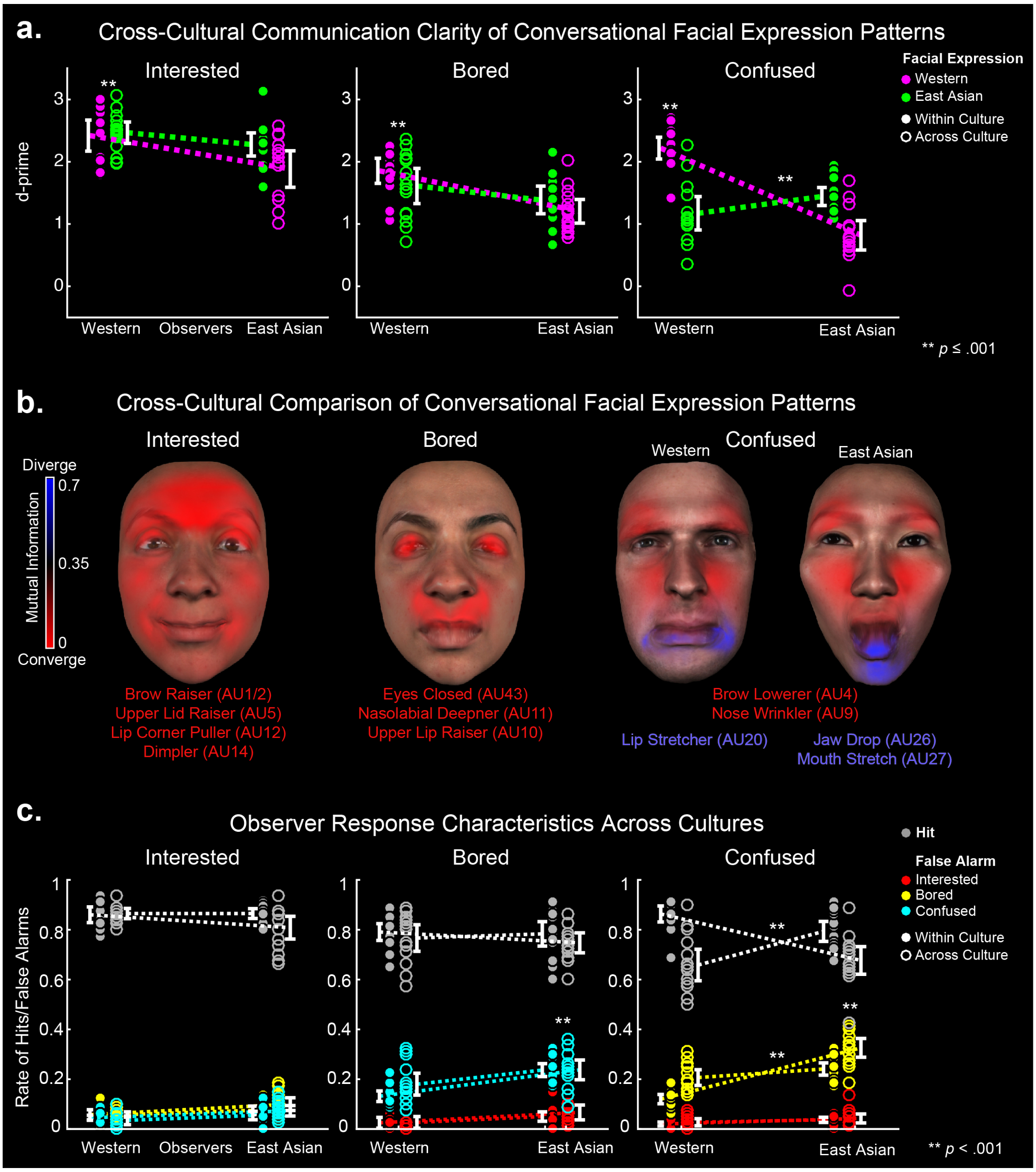

Figure 4-5. Cross-cultural communication clarity of conversational messages, cross-cultural comparison of their facial expression patterns, and observer response characteristics across cultures. (a) *Cross-cultural communication clarity of conversational facial expression patterns.* In each plot, color-coded points represent an individual facial expression pattern derived from the preliminary study (magenta – Western, green – East Asian) plotted according to its communication clarity (d-prime) within (solid circles) and across (outlined circles) cultures (see also labels on x-axis and legend on right). Dashed lines show the average differences in communication clarity across cultures. Error bars represent confidence intervals. In both cultures, facial expressions of 'interested' and 'bored' elicited comparably high clarity within and across cultures. In contrast, facial expressions of 'confused' showed a significant within-culture advantage in both cultures ($p < .001$) with significantly higher performance for Westerners than East Asians ($p < .001$) (b) *Cross-cultural comparison of conversational facial expression patterns.* Color-coded faces show for each conversational message the AUs that converge (i.e., common across cultures, low Mutual Information (MI), colored in red) or diverge across cultures (i.e., culture-specific, significantly high MI, color-coded in blue, $p < .05$, see colorbar on left). Facial expressions of 'interested' and 'bored' converge across cultures according to specific AUs (e.g., in 'bored,' Eyes Closed – AU43) whereas 'confused' shows cultural divergence characterized by specific AUs – Lip Stretcher (AU20) for Western, Mouth Stretch (AU27) for East Asian – with cultural convergence for Brow Lowerer (AU4) and Nose Wrinkler (AU9). Corresponding color-coded AU labels are listed below each face. (c) *Observer response characteristics across cultures*. For each conversational message separately, color-coded points in each plot represent the rate of hits and false alarms elicited by each facial expression pattern within (solid circles) and across cultures (outlined circles). Color

indicates the social message attributed to the facial expression (see legend on right). Dashed lines show the average difference across cultures. Error bars represent confidence intervals. In both cultures, the lower cross-cultural communication clarity for facial expressions of ‘confused’ is due to significantly lower hits ($p < .001$) and significantly higher false alarms ($p < .001$), specifically for the social message ‘bored.’ For facial expressions of ‘confused’ and ‘bored,’ East Asian observers also made significantly more false alarms for ‘bored’ and ‘confused,’ respectively than Westerners.

4.4.3 First role of culture: Structure of Conversational Facial Expression Patterns

Clear communication typically relies on sending stereotyped (i.e., low variance) structured patterns (Tinbergen, 1952). Therefore, clear cross-cultural communication should be achieved when facial expression patterns are highly similar (i.e., convergent) across cultures. Conversely, communication should break down if facial expressions are sufficiently dissimilar (i.e., divergent) across cultures. Based on the results of the cross-cultural communication task, facial expressions of ‘interested’ and ‘bored’ should each converge on stereotypical face movements across cultures, whereas facial expressions of ‘confused’ should show notable divergence on at least on a subset of face movements. In this analysis, I reveal convergent and divergent face movements (i.e., the sources of face signalling) that correspond with variance in cross-cultural communication clarity shown above. To this aim, for each conversational message separately, I used Mutual Information (MI) to identify the specific face movements (i.e., AUs) that are common across cultures and those that are culture-specific. MI simply measures in bits the relationship between culture (i.e., Western Caucasian or East Asian) and the presence/absence of each AU (e.g., Jaw Drop) in a given conversational message (e.g., confused). Thus, a high MI value indicates a strong relationship between an AU and culture – i.e., the AU is largely culture-specific; conversely, low MI values indicate a weak relationship – i.e., the AUs are common to both cultures. I applied this analysis only to AUs that appeared highly frequently ($p < .05$) across all of individual facial expression patterns for each conversational message in at least one culture.

Cross-Cultural Analysis of Conversational Facial Expression Patterns. To examine whether and how the conversational facial expression patterns converge or diverge across cultures, I identified for each conversational message separately the specific face movements (i.e., AUs) that are culturally common and those that are culture-specific. Our analysis comprised two steps applied to the 45 facial expression patterns (N = 15 per conversational message) used in the cross-cultural communication task.

Step 1: Statistically Highly Frequent AUs. For each conversational message and culture, I first identified the AUs that are statistically highly frequent using a Monte Carlo simulation method. To illustrate, consider I aim to identify the AUs that are statistically highly frequent across the Western validated 'interested' facial expression patterns. First, in the total set of 42 AU × 15 Western facial expressions of 'interested,' I computed the frequency of each AU across all of these 15 facial expressions. For example, Inner Brow Raiser (AU1) is present a total of 12 times over the 15 facial expression patterns, resulting in a 12/15 frequency. I also computed the total number of significant AUs across the facial expression patterns. For example, there are a total of 132 significant AUs across the 15 Western facial expressions of 'interested.' Next, to determine the statistical significance of each AU frequency, I used a Monte Carlo simulation method to randomly distribute (with replacement) the total number of significant AUs (here, 132 AUs) over the space of 42 AU x 15 facial expression patterns, and compute the resulting frequency of each AU. For example, on this iteration, Inner Brow Raiser (AU1) might have a frequency of 2/15. Over 1000 iterations, I therefore derived a distribution of AU frequencies, which I used to test the null hypothesis that the observed 12/15 proportion of Inner Brow Raiser (AU1) in the Western facial expressions of 'interested' is significantly higher than chance—i.e. above the 95th percentile of the randomly generated distribution of AU frequencies (i.e., one-tailed $p < .05$). If so, I call Inner Brow Raiser (AU1) a *highly frequent* AU. I repeated this procedure for each AU and for each conversational message category in each culture separately.

Step 2: Cultural Relationship of AUs. Having determined the highly frequent AUs for each conversational message category and culture in Step 1, I then examined the relationship of each AU to culture – i.e., do the highly frequent AUs distribute evenly across the facial expressions of both cultures, or are they specific to only one culture? To address this question, I quantified the relationship between each highly frequent AU and each culture using Mutual Information (MI). Low MI indicates that the two cultures share that particular AU across their facial expression patterns whereas high MI indicates that the AU is mostly present in the facial expressions of only one culture. To establish statistical significance, I used a Monte Carlo approach. For each highly frequent AU and conversational message (e.g., 'interested'), I produced a random distribution of MI values by randomly shuffling the cultural assignment of each individual facial expression pattern 1000 times, computing MI for each AU at each iteration, and then taking the maximum MI value across all AUs. I then used the distribution of maximum MI values to identify the AUs with a MI value in the 95th percentile of the distribution (Nichols & Holmes, 2002).

Figure 4-5b shows the results on color-coded face maps. For each conversational message, color-coded face maps show the AUs that converge (i.e., are common across cultures, with low MI, coloured in red) and those that diverge across cultures (i.e., are culture-specific, with high MI, coloured in blue; see colorbar on left). Corresponding color-coded AU names are listed below each face. Homogenous red colouring in 'interested' and 'bored' show that these facial expression patterns are common across cultures – for example, in 'interested' the Upper Lid Raiser (AU5, i.e., wide open eyes); in bored Eyes Closed (AU43). In contrast, in 'confused' the blue color-coding shows that these facial expressions diverge across cultures on the basis of cultural specific AUs – Westerners use Lip Stretcher (AU20) – a lateral mouth stretch – whereas East Asians use Mouth Stretch (AU27) – a vertical mouth stretch. As shown by the red color-coding in 'confused,' certain AU are also common across cultures – Brow Lowerer (AU4) and Nose Wrinkler (AU9). My results therefore demonstrate that the cross-cultural behavioural confusions elicited by facial expressions of 'confusion' are due to culture specific mouth movements that I identify here. In contrast, clear cross-cultural communication is achieved on the basis of highly convergent (i.e., culturally common) facial expressions patterns.

4.4.4 Second Role of Culture: Observer Response Characteristics

In addition to transmitting distinct patterns of information (e.g., Caldara et al., 2005; Guilford & Dawkins, 1991), clear communication relies on observers accurately detecting and discriminating the pattern, as shown by a specific response characteristics of high hit and low false alarm rates, respectively (i.e., high d-prime). In contrast, communication clarity can be compromised by either low signal detection (i.e., low hits) or low discrimination (i.e., high false alarms), or a combination of both. Therefore, examining observer response characteristics (i.e., their hit and false alarm rate) can reveal two specific sources of lower communication clarity characterized either by more conservative response strategy (i.e., a high threshold for evidence) or more liberal response strategy (i.e., low threshold for evidence; Stanislaw & Todorov, 1999), which can thus reflect which signs might be high or low risk to the observer if missed or mistaken. For example, failing to identify negative signs such as those threatening the integrity of the group (e.g., the violation of cultural norms and values) is high risk (e.g., Bradbury & Vehrencamp, 1998; Heeger, 1998; Nesse, 2005) because it can result in significant negative consequences such as conflict (e.g., Rosenblatt, Greenberg, Solomon, Pyszczynski, & Lyon, 1989). To avoid such risky situations, observers might show liberal response characteristics – that is, reporting the presence of a threat with a low threshold for evidence. Given that cultures each have a set of shared values and beliefs, examining response characteristics across cultures can therefore reveal which signs are most important to identify and provide further insights into communication systems across cultures (e.g., see Elfenbein et al., 2002 for discussion).

To further understand the sources of lower communication clarity across cultures I examined the rate of hits (i.e., correct match, 'yes' response) and false alarm responses (i.e., incorrect match, 'yes' response; see Table 4-2 for average response rates).

Table 4-2

Distribution of Hits and False Alarms for Conversational Facial Expressions in Each Culture.

		Within-Culture		Cross-Culture	
	Culture of Observer	Western	East Asian	Western	East Asian
	Culture of Facial Expression	Western	East Asian	East Asian	Western
Facial Expression	Message	*Mean (SE)*	*Mean (SE)*	*Mean (SE)*	*Mean (SE)*
Interested	**Interested**	**0.86 (0.15)**	**0.87 (0.01)**	**0.87 (0.01)**	**0.81 (0.02)**
	Bored	0.06 (0.01)	0.08 (0.01)	0.06 (0.01)	0.10 (0.01)
	Confused	0.05 (0.01)	0.06 (0.01)	0.03 (0.01)	0.07 (0.01)
Bored	**Bored**	**0.79 (0.02)**	**0.78 (0.02)**	**0.77 (0.25)**	**0.75 (0.02)**
	Interested	0.03 (0.01)	0.05 (0.01)	0.03 (0.01)	0.07 (0.01)
	Confused	0.13 (0.01)	0.24 (0.01)**	0.18 (0.15)	0.24 (0.02)**
Confused	**Confused**	**0.86 (0.02)**	**0.79 (0.02)***	**0.66 (0.03)****	**0.68 (0.03)****
	Interested	0.02 (0.01)	0.04 (0.01)	0.03 (0.01)	0.04 (0.01)
	Bored	0.12 (0.01)	0.24 (0.01)**	0.21 (0.02)**	0.33 (0.02)**

Note. *$p < .01$, ** $p < .001$

Figure 4-5c shows the results for each conversational facial expression separately. In each plot, color-coded points show the rate of hits (grey circles) and the different false alarm responses (red – 'interested,' yellow – 'bored,' and cyan – 'confused') for each individual facial expression pattern within (solid circles) and across cultures (outlined circles; see legend on right). Dashed lines show the average difference across cultures. Error bars represent confidence intervals. First, both Western and East Asian facial expressions of 'confused' showed a significantly lower hit rate across cultures than within (2-way ANOVA: 2 [culture of facial expression pattern] × 2 [culture of observer], significant 2-way interaction: F (1, 56) = 49.25, two-tailed $p < .001$, $\eta_p^2 = .47$; pairwise comparisons with Bonferroni correction, Western: two-tailed $p < .001$, Cohen's $d = 1.50$, 95% CI [0.14, 0.27]; East Asian: two-tailed $p = .001$, Cohen's $d = 1.12$, 95% CI [0.05, 0.18]). These facial expressions also elicited a significantly higher false alarm rate across than within cultures specifically for 'bored' (3-way ANOVA: 2 [culture of facial expression pattern] × 2 [culture of observer] × ['bored'/'interested' false alarms], significant 3-way interaction: F (1, 112) = 28.27, two-tailed $p < .001$, $\eta_p^2 = .50$; significant 2-way interaction for 'bored' false alarms: F (1, 56) = 40.04, two-tailed $p < .001$, $\eta_p^2 = .59$; pairwise comparisons (Bonferroni corrected) – Western: two-tailed $p < .001$, Cohen's $d = 1.37$, 95% CI [0.05, 0.13]; East Asian: two-tailed $p = .001$, Cohen's $d = 1.19$, 95% CI [0.05, 0.12]). I found no significant differences for 'interested' (F (1, 56) = 1.18, two-tailed $p = .28$, $\eta_p^2 = .04$). Thus, Western and East Asian facial expressions of 'confused' elicited both a lower hit rate and higher false alarm rate across cultures specifically for 'bored' resulting in lower cross-cultural communication clarity as show in Figure 4-5, panel (a).

A cross-cultural analysis of observers showed that for facial expressions of 'confused' (right panel) and 'bored' (central panel) East Asian observers made significantly higher false alarms than Western observers when asked to detect the social messages 'bored' (within-culture: two-tailed $p < .001$, Cohen's $d = 1.65$, 95% CI [0.07, 0.17]; cross-culture: two-tailed $p < .001$, Cohen's $d = 1.93$, 95% CI [0.07, 0.12]) and 'confused' (main effect of culture of observer: F (1, 56) = 26.36, two-tailed $p < .001$, $\eta_p^2 = .49$), respectively. I found no such differences for facial expressions of 'interested' (F (1, 56) = 1.85, two-tailed $p = .18$, $\eta_p^2 = .06$). Together, these data suggest that East Asian observers use a more liberal response strategy for negative social messages – that is, they are more likely to report 'signal present' when asked to detect the social messages 'confused' and 'bored' – than Westerners when viewing negative facial expressions. While producing more false alarms (i.e., detection errors), this type of response characteristics could be adaptive if there are high risks associated with missing negative social messages within the socio-cultural environment, such as threatening the social harmony of the group (e.g., Hofstede & Hofstede, 2001; Markus & Kitayama, 1991; Matsumoto, 2006).

4.5 Discussion

Using subjective cultural perception, a facial expression generator, and reverse correlation, I objectively characterized the dynamic facial expression patterns that communicate key conversational messages within Western and East Asian culture. In a cross-cultural communication task, I showed that Western and East Asian facial expressions of ‘interested’ and ‘bored’ elicited clear communication across cultures. In contrast, facial expressions of ‘confused’ diminished cross-cultural communication clarity. To understand the sources of variable cross-cultural communication clarity I analysed two key components of communication in each culture: the structure of facial expressions that transmit these social messages, and the response characteristics of the observers. Using a novel information theoretic analysis applied to the facial expression patterns, I show that facial expressions of ‘confused’ diverge across cultures based on specific mouth movements, thereby hindering cross-cultural communication of ‘confusion.’ In contrast, facial expressions of ‘interested’ and ‘bored’ are understood clearly across cultures because they converge across cultures. Analysis of observer response characteristics also showed that East Asians are more likely to incorrectly report the presence of the negative social messages ‘bored’ and ‘confused’ when viewing other negative facial expressions (i.e., ‘confused’ and ‘bored,’ respectively). Such a strategy means that East Asian observers are more likely to avoid risks associated with missing these negative social messages, such as conflict (e.g., Hofstede & Hofstede, 2001; Markus & Kitayama, 1991; Matsumoto, 2006) while minimizing the risks associated with incorrect detection (i.e., false alarms) because they occur in the context of other negative social signals.

Facial expressions of ‘confused’ also showed cultural commonalities including Brow Lowerer (AU4) and Nose Wrinkler (AU9), each of which is known to modulate sensory exposure by diminishing visual and olfactory input, respectively (Susskind et al., 2008) and could indicate rejection or an obstacle (Scherer, Mortillaro, & Mehu, 2013). Such physiological functions are commensurate with the social function of terminating or re-routing an interaction, which, like ‘interest’ and ‘bored,’ suggests that facial expressions of ‘confusion’ have a physiological origin. Facial expressions of ‘confused’ also showed clear cultural differences in mouth movements. Specifically, Western patterns showed a lateral mouth stretch (Lip Stretcher; AU20), whereas East Asians use a vertical mouth stretch (Mouth Stretch; AU27), as reflected by Western (:-/) and East Asian (+o+) emoticons. These culturally divergent facial expressions did not elicit clear communication across cultures, suggesting that facial expressions of confusion might be learned within a culture. For example, Lip Stretcher (AU20) and Mouth Stretch (AU27) are observed in facial expressions of pain in adults and infants across Western and East Asian cultures (Williams, 2002), with wider mouth opening associated with more intense pain (Hicks et al., 2001; Li et al., 2007). Thus, while the message ‘confusion’ could comprise an element of (social) pain in each culture, the magnitude might differ across cultures and cause cross-cultural confusions. My data therefore suggests that across cultures, facial expressions of ‘confusion’ share a common biological origin, but with distinct culture-specific accents that are learned (e.g., Elfenbein, 2013; Jack et al., 2016; Jenssen, 1979).

Finally, analysis of observer response characteristics showed that East Asian observers are more likely to incorrectly report the presence of the negative social messages ‘bored’ and ‘confused’ than Westerners when viewing other negative facial expressions (i.e., ‘confused’ and ‘bored,’ respectively). My data suggests that East Asians require less evidence to report detecting negative social messages than Westerners such as when viewing another negative (but not necessarily physically similar) signal. My results correspond with previous facial expression recognition studies showing that observers in collectivist cultures tend to make more errors for negative facial expressions of emotion than individualistic culture (e.g., see Matsumoto, 2006 for a review). However, my results also contradict previous interpretations that collectivist cultures are *less* likely to report the presence of negative social signals than individualistic cultures to protect social harmony (e.g., Matsumoto, 1989, 1992). Rather, in line with signal detection theory (e.g., Heeger, 1998) and fundamental theories of communication (e.g., Bradbury & Vehrencamp, 1998; Nesse, 2005) since collectivist cultures place a higher value on social harmony than individualistic cultures (e.g., Hofstede, 1984; see Markus & Kitayama, 1991 for a review), East Asian observers should be more likely to report the presence of a negative social signal with less evidence, such as another negative signal (see also e.g., Schultz, Izard, & Ackerman, 2000) in order to avoid the high risks associated with missing the signal.

In sum, I show that culture plays a dual role in impacting communication clarity within and across cultures by shaping both the signalling structure of conversational facial expressions and observer response characteristics. My data are consistent with the biological origins of facial expressions (e.g., Darwin, 1872/1998; Susskind et al., 2008) while still supporting theories of cultural learning (e.g., Barrett et al., 2011; Elfenbein, 2013). My results therefore have direct implications for cross-cultural communication, including the design of digital agents that must automatically detect and produce culturally appropriate facial expressions (e.g., Breazeal, 2003; Cassell et al., 2000; De Rosis et al., 2003; Foster et al., 2012; Gratch, DeVault, Lucas, & Marsella, 2015; Gratch et al., 2002).

5 General Discussion

In this thesis, I first investigated how facial movement patterns are used to communicate a broad set of social messages in Western and East Asian culture. Specifically, I first examined the facial expressions of diametrically opposite concepts of physical pain and pleasure in Study 1. My results demonstrated the physical and perceptual distinctiveness of facial expressions of pain and pleasure and therefore each can be used to communicate efficiently the corresponding message in social communication. This challenges the previous studies that the facial expressions of pain and pleasure are highly similar and communicating such messages should rely on the other sources of information such as the body. By further examining the cultural commonalities and specificities of facial expressions of pain and pleasure, I show that across cultures pain facial expressions are highly similar. In contrast, facial expressions of pleasure formed distinct culture-specific clusters, suggesting the role of cultural learning and their function as evolved social signals. Thus, my results also highlight the potential role of biological and cultural factors in shaping facial expressions of pain and pleasure.

To further understand the relationship between facial movement patterns and their corresponding social messages, I then expanded my research to a broad set of mental states in Western and East Asian culture in Study 2. The NMF analysis reveals four culturally common facial movement patterns that may reflect their physiological origins – for example, to regulate the sensory input by opening or closing the eyes, using nose wrinkling to show rejection. Interestingly, whereas emotional facial expressions primarily reply on such fundamental facial movement patterns, conversational facial expressions show more complex facial movements that are not presented in the fundamental facial movement patterns. These facial movements may have evolved for social communication among cultural individuals and therefore diverge across cultures. These results motivated me to conduct a cross-cultural study to further understand whether such conversational facial expressions converge and diverge across cultures and how this can impact cross-cultural communication.

Thus, to further understand the role of culture in impacting the signaling and receiving of facial expressions, I conducted a cross-cultural study using the four key conversational facial expressions in Study 3. Specifically, I first objectively characterized the structure of dynamic facial expression patterns that communicate key conversational messages – ‘interested,’ ‘bored’ and ‘confused’ – in Western and East Asian culture. Next, in a cross-cultural communication task, I show that facial expressions of 'interested' and 'bored' are understood across both cultures whereas 'confused' causes cross-cultural confusions. Using a novel information-theoretic analysis called Mutual Information, I show that these cross-cultural confusions are due to culture-specific mouth movements in 'confused,' whereas 'interested' and 'bored' involve culturally common, biologically rooted face movements (e.g., eye whites). Finally, I examined the observer response characteristics and show that cultural specificities that also impact communication clarity. That is, East Asian observers more often report the presence of negative messages (i.e., 'confused' and 'bored') than Westerners when viewing other negative expressions.

Together, my data reveal for the first time how specific facial movement patterns are used to communicate a broad set of social messages including physical states and mental states in Western and East Asian cultures. More specifically, my results mapped the facial expressions with the semantic network of their corresponding social messages and demonstrated in details how semantic similar/distinct social messages are delivered by similar/distinct facial movement patterns – a theory that was proposed by Darwin (1872/1998). Analysis of a broad set of facial expressions in two cultures reveals the different structure of facial expressions, with the fundamental movement patterns that are primarily used to communicate emotional messages and more complex movement patterns that are used for delivering other messages such as manipulating the flow of conversations. These different components of facial expressions may reflect the physiological origins and communication functions as they were developed in the evolution. Finally, my results show for the first time how culture impacts the both sides of social communication – signalling and receiving facial expressions. This reveals the source of confusion in cross-cultural communication, which has direct implications for understanding cross-cultural social communication, including the design of socially interactive digital agents. Here, I will further discuss each point with links to my future work.

5.1 The Principle of Antithesis: Facial Movement Patterns Reflect the Structure of Corresponding Social Messages

The representation of a given social signal can inform the nature and form of non-verbal communication. According to Darwin's principle of antithesis (Darwin, 1872/1998), distinct facial movement patterns have been evolved to communicate distinct social messages such as 'pain' and 'pleasure,' 'interested' and 'bored,' which often reflects their physiological function and real world iconicity. For example, widening the eyes and raising the eyebrows increases visual input and could therefore be used to show interest and a desire for more information (e.g., Darwin, 1872/1998; Izard, 2013), whereas lowering eyebrows and closing eyes decrease the sensational input and therefore typically to show the social signal of rejection (e.g., Lee et al., 2014; Susskind et al., 2008). My data support this view by showing that ‘pain’ and ‘pleasure’, ‘interested’ and ‘bored’ each comprise culturally common face movements that are known to modulate sensory exposure in opposite ways (see also discussion by Hass, 1970; Izard, 2013; Tomkins, 1984). Furthermore, my data reveals that facial expressions for communicating a board set of mental states rely on four fundamental facial movement patterns, which are consistently associated with valence and arousal across cultures. For example, facial expressions of positive mental states primarily show smiling, whereas facial expressions of negative mental states show either stretching lips (i.e., the opposite lip movement as for smiling) or nose wrinkling that originated from avoiding bad smell and taste (e.g., Chapman et al., 2009). Thus in line with biological signalling accounts that predict distinct messages are communicated using distinct information patterns (e.g., Caldara et al., 2005; Guilford & Dawkins, 1991), my results demonstrated that the distinct social messages of are each communicated using distinct and iconic face movement patterns with their physiological origins.

5.2 The Evolution of Facial Expressions: Physiological Origin and/or Communication Function

My results show that facial expressions of social messages such as ‘pain,’ ‘interested’ and ‘bored’ converge across cultures, reflecting a shared knowledge of these facial expressions. According to Darwin’s account of the biological origins of facial expressions (Darwin, 1872/1998), such convergence could arise due to socially functional facial expressions retaining face movements that once served a physiological function. My data support this view by showing that ‘pain,’ ‘interested’ and ‘bored’ each comprise culturally common face movements that also modulate sensory exposure (Susskind et al., 2008). Specifically, ‘interested’ involves the Upper Lid Raiser (AU5) and Inner and Outer Brow Raiser (AU1-2), which increase the intake of visual information. Similarly, ‘bored’ involves Eyes Closed (AU43) and Upper Lip Raiser (AU10), which diminish visual and olfactory sensation, respectively. The physiological role of each facial movement is also consistent with their social functions. For example, communicating ‘interest’ by showing an increase in sensory intake implies novelty (see Scherer et al., 2013 for a review) and a desire for continued input, thereby encouraging the interaction to be extended (Chovil, 1991). Communicating ‘bored’ by showing a reduction in sensory intake could indicate rejection (e.g., 'I don't want to see it'; Bavelas & Chovil, 1997) and terminate or re-route the interaction.

5.3 The Dual Role of Culture on Signalling and Receiving Facial Expressions

Using the data-driven approach of reverse correlation and signal detection theory, my results provide a novel paradigm to understand how culture impacts cross-cultural communication by examining the two main components of a communication system – the sender and the receiver. For the senders, I characterized the culture specific facial expressions in social interaction including physical states (e.g., ‘orgasm’) and conversational messages (i.e., ‘confused’). I applied Mutual Information analysis and objectively identified the culture-specific facial movement patterns, which is typically difficult to achieve using theory-driven approaches (see reviews of Barrett et al., 2011; Elfenbein, 2013; Russell, 1994). I further examined how such cultural specificities impact cross-cultural communication in Study 3. Cross-cultural confusion on the facial expression of ‘confused’ confirmed the validation of using such paradigm to characterize the culture specificities of facial expressions and suggesting the role of cultural learning in the evolution of facial expression (e.g., Barrett & Gendron, 2016; Elfenbein, 2013; Shariff & Tracy, 2011).

For the receivers, I applied signal detection approach and examined the observer response characteristics, which reveals that East Asian observers are more liberate on detecting negative social messages. That is, East Asians require less evidence to report detecting negative social messages than Westerners such as when viewing another negative signal. This cultural difference of response characteristics may also reflect in Study 2, where East Asians tend to perceive high arousal facial expressions as negative and Westerns perceived a mix of positive and negative mental states for such facial expressions. Thus, it is possible that East Asians are more likely to detect negative information from high arousal facial expressions. Together, my results suggest that East Asians are more sensitive to detect negative social messages from others, which may reflect the close relationship between self and others in the collectivist culture (see a review of Markus & Kitayama, 1991).

5.4 Other Sources of Non-Verbal Communication

Finally, understanding how facial expressions converge and diverge across cultures can also inform the nature and form of non-verbal communication. For example, although 'interested,' 'bored,' and 'confused' could comprise multimodal signals involving vocalizations (e.g., Schegloff, 1982) and body postures their facial expressions could be understood clearly within a culture at typical social interaction distances (e.g., Smith & Schyns, 2009). In contrast, facial expressions of 'thinking' did not elicit clear communication with culture suggesting that this message relies on other sources of multi-modal information such as vocalizations (e.g., Chovil, 1991), averted gaze (e.g., Harness Goodwin & Goodwin, 1986), head movements (e.g., Cunningham, Breidt, Kleiner, Wallraven, & Bülthoff, 2003), and/or hand gestures (e.g., Mahmoud & Robinson, 2011). Such sources of information will be the focus of future research to understand how they each contribute (either independently or in combination) to social communication within and across cultures.

In sum, by objectively characterizing the structure of dynamic facial expression patterns and examining the signaling and receiving of such facial expressions, my results reveal for the first time how specific facial movement patterns are used to communicate a broad set of social messages in Western and East Asian culture and how culture shapes the signalling and perception of facial expressions in cross-cultural communication. Next, I will discuss the implication of my results in the field of psychology, computer science and social robotics, with links to my future work on developing a mathematical model of face social signalling and transfer this knowledge to socially and culturally sensitive conversational agents.

5.5 Future work

5.5.1 Examining Cross-cultural Communication Using a Broad Set of Dynamic Facial Expressions

The twenty-first century witnessed the rapid globalization and cultural integration. The advances of transportation (e.g., jet engines, containerships) and telecommunication technology (e.g., Internet, mobile phones) have transcended the geographical constraints and changed the way that human beings communicate. For example, social media has become an essential part for our daily life (see Kaplan & Haenlein, 2010 for a review). Over 10 billion photographs and videos are uploaded to the online hosting sites such as

Flickr and YouTube, which makes the total collection of 8 million objects in the British Museum looks like such a tiny piece of cake. However, due to ‘a smaller world’ that bring all cultural groups together, the globalization and the rising digital economy (e.g., as defined by Tapscott, 1996) increase the pressure of cross-cultural communication (see a review of Ting-Toomey, 1994). Cultural groups can differ a lot in their ways of communicating (Archer, 1997; Elfenbein et al., 2007) and successful cross-cultural communication is not always guaranteed. For example, according to the culture values that are measured in Hofstede et al. (1990), individuals in a culture with high power distance (i.e., a culture that accepts the power is distributed unequally among the members) are more likely to legitimize the aggressive behaviours (e.g., the facial expression of anger or disgust) from the upper class. In contrast, individuals in a culture with low power distance are more likely to perceive these behaviours as an inappropriate and unequal treatment (Bond, Wan, Leung, & Giacalone, 1985; Tedeschi, Gaes, & Rivera, 1977).

Consequently, it is crucial to understand how a broad set of facial expressions are used by cultural individuals and what elicits the confusion in cross-cultural communication. This knowledge will not only contribute to further understanding the function of facial expressions as social signals in different cultures, but also providing insights into how culture shapes the way that people interact with each other. The practical implication of this knowledge will certainly benefit the individuals by being aware of such confusions in cross-cultural communication and therefore improve their communication skills.

In this thesis, I have examined the cross-cultural communication using the four facial expressions and show that culture impacts both how the sender used the social signals (i.e., facial movement patterns) to communicate social message and how the receiver perceived such social message (i.e., the observer characteristics). In my further work, I would like to further examine the role of culture in impacting social communication by expanding to a board set of facial expressions. Specifically, I would like to explore the ‘complex’ facial expressions including conversational facial expressions, which were probably evolved for their function of social communication (Shariff & Tracy, 2011) and therefore tuned by cultural context (Darwin, 1872/1998). Apart from the Action Unit patterns that is the main focus of this thesis, I would like to explore the cultural convergence/divergence in other aspects of dynamic facial expressions – for example, amplitude, early or late peak latency (e.g., Jack et al., 2014).

5.5.2 Equipping Social Interactive Agents With Socially and Cultural Sensitively Dynamic Facial Expressions

When the world chess champion Garry Kasparov was beaten by Deep Blue (a chess machine built by IBM) in the human-computer chess match in 1997, the rapid development of artificial intelligence triggered a worldwide debate on the future of artificial intelligence. Researchers such as Hans Moravec and Ray Kurzweil celebrated it as the beginning of a beautiful new era (i.e., technological singularity), where humans can build such machines that transcend the biological constrains and transform the human intelligence into an immortal digital form (Kurzweil, 2005). Others such as Steven Hawking and Elon Musk warned that the advanced artificial intelligence (i.e., "superintelligence" in Bostrom, 2014) “could spell the end of the human race” if it goes beyond human control (Cellan-Jones, 2014). No matter if the technological singularity will be animated by love or by wrath in the future, both groups of researchers agree that the current development of human-level artificial intelligence (e.g., self-driving cars, automatic medical diagnosis, personal assistants such as Siri) is useful to many aspects of human society.

However, even it can beat the world chess champion, a cold flickering screen seems far away from ‘communication.’ Over the past two decades, computer scientists have developed a number of robots to serve different social function such as security (e.g., Song, Yin, Zhou, & Cheng, 2009), healthcare (e.g., Broadbent et al., 2010), personal assistant (e.g., Dario, Guglielmelli, Laschi, & Teti, 1999), pet (e.g., Kubota, Nojima, Baba, Kojima, & Fukuda, 2000), retailing (Kamei et al., 2010), teaching (e.g., You, Shen, Chang, Liu, & Chen, 2006) and entertaining (Andersson, 1989). In the similar vein, virtual humans have also been developed using the computer-generated images and sound (e.g., Becker-Asano & Wachsmuth, 2010; Khullar & Badler, 2001; Rickel et al., 2002), with the aim of simulating the social communication in virtual reality (Steuer, 1992). But very few of such robots or virtual agents are actually perceived as ‘social’ by human participants and even fewer are actually applied to commercial use as social interactive agents (Hegel, Lohse, & Wrede, 2009).

To engage their human users, social robots must be equipped with knowledge of communication behaviours including recognizing and/or producing a broad set of face expressions to communicate with their human users (e.g., Cassell et al., 2000; Poggi & Pelachaud, 2000). One famous example of how it can go wrong is the *uncanny valley* –

when a robot looks very much like a human from its appearance (one recent example is a human-like robot called Sophia from Hanson Robotics) but does not behave as 'human-like.' Human observers often show a strong revulsion and refuse to interact with such a robot (Mori, MacDorman, & Kageki, 2012). As a result, more and more computer scientists in the field start to realize the importance of applying the knowledge of human communication in their design. For example, in the book of *Embodied Conversational Agents* (Cassell et al., 2000), researchers argue that facial expressions that supporting the conversations (both emotions and the other mental states) are essential for a digital agent to be able to communicate with human – to pass a face-to-face Turing test. The lack of such a system that can encode and decode such facial expressions (e.g., showing confused, thinking or agreement) will make the digital agent perceived as not able to communicate properly and unnatural (Cassell, Bickmore, Campbell, Vilhjalmsson, & Yan, 2001). Therefore, Takeuchi and Nagao (1993) concluded that facial expressions of mental states are a new modality to coordinate the human-computer interaction. Moreover, since the rules of communication can change across cultures, it is even more challenging to develop culture-aware social robotics.

To design and develop social interactive agents, social roboticists often turn to the field of Psychology – particularly social psychology and vision science – to understand what it is about the face that's driving these social judgments in humans. However, the face is incredibly complex – it can make hundreds of facial expressions, it varies in its shape in innumerable ways, and comes in a broad variety of colours and textures, which then makes finding what it is in the face that drives our social judgments a genuine empirical challenge (e.g., see a review of Jack & Schyns, 2017). This knowledge gap then prevents the role of knowledge of facial expressions in human-human communication in designing and developing human-computer interactions.

By mathematically modelling and validating a broad set of facial expressions used in Western and East Asian culture, my data can address this knowledge gap by building a generative system of facial expressions and transferring the knowledge to social interactive agents. Specifically, using this generative grammar of face movements, digital agents can easily produce a wide variety of facial expressions by adding different accents to these core facial movement patterns. This flexibility will then allow the possibility of using social agents in a variety of social and cultural scenarios, which is one of the main goals in the field (e.g., Gockley et al., 2005; Gratch et al., 2002; Zhang et al., 2007). More importantly, the variability of communication behaviours, for example, different ways of smiling, is a key component for improving the naturalness of a social interactive agent (e.g., Hiatt, Harrison, & Trafton, 2011). The facial expression patterns in my data are developed from each cultural individual and therefore characterized the natural individual differences in the population. Social interactive agents that ‘personalized’ by these natural differences have direct implications on not only producing natural and realistic communication behaviours, but also designing personal assistive robots that can be precisely tuned for their human user (e.g., Breazeal, 2004).

5.5.3 Developing a Mathematical Model of Social Face Signaling: Dynamic Facial Expression, Morphology and Complexion

Neuroscience studies have suggested that human brain evolved two relatively independent system for processing facial expressions and face identity (e.g., Calder & Young, 2005; Haxby, Hoffman, & Gobbini, 2000). Thus, apart from dynamic facial expressions, other components including morphology and complexion that determines the identity of a face also play an important role in driving different social judgments. Previous studies have shown even from a brief glance at a static image of a face or part of it, observers can make inferences for age (e.g., Fink, Grammer, & Matts, 2006; Van Rijsbergen et al., 2014), gender(e.g., Smith et al., 2005), health(e.g., Fink & Matts, 2008), attractiveness (e.g., Jones, Little, Burt, & Perrett, 2004; Scheib, Gangestad, & Thornhill, 1999), trustworthiness(e.g., Oosterhof & Todorov, 2008; Todorov, Baron, & Oosterhof, 2008) and dominance (e.g., Gill et al., 2014). Thus, the next step is to combine dynamic facial expressions with the other main components of the face –morphology and complexion to develop a comprehensive model of social face signalling. This comprehensive model can then allow the possibility for the researchers in psychology and neuroscience to manipulate the involuntary and/or voluntary facial signals (e.g., Adolphs, 2002; Calder & Young, 2005; Haxby et al., 2000; Smith et al., 2005; Todorov, Said, Engell, & Oosterhof, 2008), which then provides powerful approaches to precisely identify their functions in social communication. By transferring this powerful social face signaling to social interactive agents in the future, my work will also benefit the field of computer science by flexibly generating and combining dynamic facial expressions and face identities for a variety of social settings.

5.5.4 Combining the System of Facial Signals With Other Sources of Multi-Modal Communication

5.5.4.1 Combining With Other Sources of Non-Verbal Communication

Human beings are masters in using not facial expressions but also other nonverbal signals such as speech-independent gestures and body movements to communicate. A ‘V-for-victory’ gesture or shaking fists can quickly change the judgment of other individual’s social judgments and therefore are curial in daily life. Nonverbal signals are pervasive in our day-to-day communication including informal chat and formal job interviews. Such nonverbal communication is a crucial component for evaluating individual’s social traits (e.g., trustworthiness, dominance, leadership), job-related qualities and performance in personnel selection, business application and politic relations (e.g., Eagly, 1995; Gifford, Ng, & Wilkinson, 1985; Granhag & Strömwall, 2002; Riggio, 2005). In addition, the sex-related difference in nonverbal behaviors can produce different preferred properties of mates, attitudes to sexual behavior (see Eagly, 1995 for a review) or even misinterpretation of sexual provocation in the workplace (Woodzicka & LaFrance, 2005).

Some nonverbal behaviors such as gestures and touch are encoded with conscious intention and control (see Knapp et al., 2013 for a review). Other nonverbal signals such as postural expansion (i.e., showing dominant, see Tiedens & Fragale, 2003) and postural constriction (i.e., showing submissive) can be processed automatically by individuals even without realizing them (see Lakin, 2006 for a review). An interesting example of how nonverbal behaviors can affect the other individual’s behavior is the *back channel*s of the listener’s behaviors during a conversation. The listeners’ responses such as facial expressions (e.g., smiling or looking suspicious), head movement (e.g., nodding or shaking head), gaze (direct eye contact or looking somewhere else) and generic nonverbal vocalizations (e.g., “mhm” or “uh-huh”) can almost certainly and immediately change the behavior of the speaker during the spontaneous conversations (see a review of Bavelas & Chovil, 2000). Even when the speaker is absolutely dominant (i.e., the listener is not allowed to speak during the experimental control condition), the conversation still operates as “a duet, not a solo” (Clark, 1996). Bavelas et al. (2000) show that if the listeners show lack of simultaneous back channel behaviours of understanding (which was modulated by increasing their cognitive load by doing a counting task), the storytellers (i.e., dominant speakers) will become more influent in speaking as they are waiting for “some kind of reactions,” make the story significantly shorter and report much more unpleasant experience after the experiment.

My results show that whereas a board set of social messages can be communicated via facial expressions, some mental states (e.g., 'thinking') require other sources of non-verbal communication. Thus, in my future work, I would like to combine the dynamic facial expressions with other forms of non-verbal communication such as gaze, hand and body gestures and examine the role of each social signal in the multi-modal communication.

5.5.4.2 Combining With Verbal Communication

One of the crucial distinctions between human beings and any other organisms on the planet is language – the ability to learn and use the complex communication systems of acoustic and visual symbols. Although several case studies report that animals such as parrot (Pepperberg, 1981), border collie (Pepperberg, 1981) and bottlenose dolphin (Herman, Richards, & Wolz, 1984) can understand simple words and sentences, there is a unique boundary between the animals that can use a limited set of words and syntax (most likely after thousands times of reward and punishment) and a human child who can infer the meaning of the word from its context, adopt the grammar structure and apply the rules to generate infinite brand new sentences that he or she never came across before (e.g., Bruner, 1985).

Verbal and non-verbal signals both play a central role in manipulation daily social interaction. For example, turning-taking point is when the individuals co-ordinate and exchange who is to speak next. To avoid any unnatural gap or overlapping in the conversation, speakers have to signal at the right time of turn transition via *verbal signals* such as grammar prosody and pragmatics in English (e.g., De Ruiter, Mitterer, & Enfield, 2006; Sacks et al., 1974; Stephens & Beattie, 1986), sentence-final marking words in Korean and Japanese (e.g., Kim, 1999; Tanaka, 2000), and *non-verbal signals* such as a gaze from the speaker to the listener (e.g., Duncan & Fiske, 2015; Sacks et al., 1974). Stivers et al. (2009) analyzed the conversations in 10 languages and show that the time sequences in a turning-talking system is overall universal in different languages and in most cultures the gap is within 200 milliseconds before the next speaker starts his turn of speaking. However, speakers from different cultures do have slight different tolerance of the delay and the overlaps of exchanging the turn. Italian speakers are more used to overlapping-talk, whereas Japanese speakers tend to leave a noticeable gap of silence before the next speaker's turn. Even when a conversation does not going that well, speakers use words such as ‘uh’ or ‘um[2]’ as ‘fillers’ to fill the minor or major pause and make the speech more fluent. Clark and Tree (2002) argue that the function of these words is not to indicate a problem (i.e., performance error in speaking) as often suggested in the literature (e.g., Chomsky, 2014; Eisler, 1968). Instead, ‘uh’ and ‘um’ are used as a predictor for upcoming pause and delay before speaking, which is an essential component for the on-going performance of a conversation system. Unlike treating them as errors, speakers process these fillers in the same way as for other meaningful words.

Most nonverbal behaviors are encoded consistent with verbal behaviors. They either operate to repeat what was said or to assist the understanding of the meaning. For example, individuals often raise or lower eyebrows simultaneously with the speech as a syntactic marker to emphasize a word (i.e., *underliner*), suggest a pause (i.e., *punctuation*) or ask a question (i.e., *question maker*) in spontaneous conversation (e.g., Bavelas & Chovil, 1997; Chovil, 1991; Ekman, 1979). *Co-speech gestures* (also referred as *illustrators*) such as drawing the shape, direction and movement of an object can help to describe it by visualizing the corresponding properties. It is not uncommon, however, that nonverbal behaviors are incongruous with what the individual says. An everyday example of such inconsistent message is *sarcasm* – when the nonverbal channel (e.g., pitch, tone) intentionally contains the contradictory affect against the meaning of words (e.g., Mehrabian & Wiener, 1967). Sometimes such incongruent behavior can result in unpleasant or harmful consequences to other individuals, especially for the young ones who may get confused and interpret it negatively. Bugental, Love, Kaswan, and April (1971) found that a positive correlation between the disturbed children (i.e., who often demonstrate behavioral or emotional disturbance in the classroom and playground) and the conflicting message between verbal and nonverbal behaviors of their parents. A number of studies argue that a) nonverbal behaviors are more salient than verbal cues (e.g., Argyle et al., 1970) and b) individuals tend to trust the nonverbal over verbal signals because the nonverbal behaviors (e.g., flush, Duchene smile and eye behavior) are more difficult or even impossible to manipulate and fake. Mehrabian and Wiener (1967) demonstrated that the verbal channel interprets less than 10% of the total impact of an emotion, whereas over 85% is attribute from either the facial expression or the vocal paralanguage (e.g., pitch, tone). Similarly, in their Social Interpretations Task (i.e., making judgments on a video tape comprising 20 sequences of spontaneous behaviours), Archer and Akert (1977) show that individuals make more accurate judgments of the behavior with the nonverbal cues

[2] It is usually spelled as *er* and *um* in British English (Cassell et al., 2001).

than verbal cues. However, other researchers suggest that the preference of relying on verbal or non-verbal behaviors varies across the individuals who differ in the ability of interpreting the signals from the two channels and the intention to which information they want to understand. Individuals who are more skilled with the language tend to rely on the verbal cues much more than the ones who are not native speakers. On the other side, Hall and Schmid Mast (2007) concluded that nonverbal behaviors are more often used to interpret someone's feeling, whereas the verbal behaviors are attributed toward someone's thinking. In most case, verbal and non-verbal signals cooperate, rather than compete, to achieve the same goal of communication.

Thus, the ultimate goal of my research in the near future is to model the comprehensive system of human social communication across cultures including non-verbal and verbal signals. I would like to then transfer this knowledge of human-human social communication to develop conversational agents, which will in turn assist in understanding the fundamental components of human social communication.

6 Conclusion

Human survival critically relies on communicating a broad set of social messages including physical states and mental states. The prerequisite for any successful social communication is the shared knowledge between the sender and the receiver about what and how a specific social signal is used. To communicate the broad set of social messages in daily life, human beings have developed complex facial movement patterns as one of the most important and powerful social signals. With increasing globalization, cross-cultural interactions are fast becoming integral to modern life, which presents increasing pressure for cross-cultural communication. Specifically, a broad set of facial expressions including conversational facial expressions is critical for clear communication because they guide the flow of social exchanges. Yet, our knowledge of such facial movement patterns is relatively limited in terms of their functions in different cultural context – for example, whether these important everyday facial expressions are understood across cultures or cause cross-cultural confusions.

In this thesis, I explored how facial movement patterns are used in Western and East Asian culture to communicate a broad set of social messages including physical states and mental states. Specifically, I conducted three studies where I used 4D computer graphics and the social psychophysics approach to objectively characterize the structure of dynamic facial expression patterns that communicate a broad set of social messages including physical states and mental states. I also examined the role of culture in impacting cross-cultural communication clarity using the four key conversational facial expressions. Together, my results reveal for the first time how specific facial movement patterns are used to communicate a broad set of social messages in Western and East Asian culture and how culture shapes the signalling and perception of facial expressions in cross-cultural communication. My results have direct implication in the field of psychology, computer science and social robotics by building the bridge between the knowledge of human-human social communication and human-robot communication.

Based on the results of this thesis, my future work will aim to develop a comprehensive mathematical system of face social signalling in Western and East Asian cultures by combining dynamic facial expressions, morphology and complexion of faces. I will also transfer this knowledge to socially and culturally sensitive conversational agents and contribute to improving the design and develop of socially interactive robots and virtual humans.

Appendix A

Screening Questionnaire

Western observers

Each potential observer completed the following questionnaire. I selected only those answering 'no' to all questions for participation in the experiment.

1. Have you ever lived in non-Western* country before (e.g., on a gap year, summer work, move due parental employment)?
2. How many weeks have you spent in a non-Western country (e.g., on vacation)?
3. Have you ever dated or had a very close friendship with a non-Western person?
4. Have you ever been involved with any non-Western culture societies/groups?

*by Western groups/countries, I are referring to Europe (East and West), USA, Canada, United Kingdom, Australia and New Zealand.

East Asian observers

Each potential observer completed the following questionnaire. I selected only those answering 'no' to all questions for participation in the experiment.

1. Have you ever lived in non-East Asian* country before (e.g., on a gap year, summer work, move due parental employment)?
2. How many weeks have you spent in a non-East Asian* country (e.g., on vacation)?
3. Have you ever dated or had a very close friendship with a non-East Asian* person?
4. Have you ever been involved with any non-East Asian* culture societies/groups?

*by East Asian groups/countries, I are referring to China, Japan, Korea, Thailand and Taiwan.

References

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

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature, 372*(6507), 669.

Ahumada, A., & Lovell, J. (1971). Stimulus features in signal detection. *The Journal of the Acoustical Society of America, 49*(6B), 1751-1756.

Ahumada, A. J. (2002). Classification image weights and internal noise level estimation. *Journal of Vision, 2*(1), 8-8.

Alberti, L., & Kemp, M. (1435/1991). *On painting* (Vol. 21): Penguin UK.

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological monographs, 47*(1), i.

Anand, K., Rovnaghi, C., Walden, M., & Churchill, J. (1999). *Consciousness, behavior, and clinical impact of the definition of pain.* Paper presented at the Pain Forum.

Anderson, A. K., Christoff, K., Panitz, D., De Rosa, E., & Gabrieli, J. D. (2003). Neural correlates of the automatic processing of threat facial signals. *Journal of Neuroscience, 23*(13), 5627-5633.

Andersson, R. (1989). Dynamic sensing in a ping-pong playing robot. *IEEE Transactions on Robotics and Automation, 5*(6), 728-739.

Archer, D. (1997). Unspoken diversity: Cultural differences in gestures. *Qualitative Sociology, 20*(1), 79-105.

Archer, D., & Akert, R. M. (1977). Words and everything else: Verbal and nonverbal cues in social interpretation. *Journal of personality and social psychology, 35*(6), 443.

Argyle, M., Salter, V., Nicholson, H., Williams, M., & Burgess, P. (1970). The Communication of Inferior and Superior Attitudes by Verbal and Non-verbal Signals. *British journal of social and clinical psychology, 9*(3), 222-231.

Aristotle. (1913/1999). Physiognomonica.

Aristotle. (2000). *On the soul*: Alex Catalogue.

Aristotle, J. B. (1885/1984). *The complete works of Aristotle*: Princeton, NJ: Princeton University Press.

Aviezer, H., Messinger, D. S., Zangvil, S., Mattson, W. I., Gangi, D. N., & Todorov, A. (2015). Thrill of victory or agony of defeat? Perceivers fail to utilize information in facial movements. *Emotion, 15*(6), 791.

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science, 338*(6111), 1225-1229.

Back, E., Jordan, T. R., & Thomas, S. M. (2009). The recognition of mental states from dynamic and static facial expressions. *Visual Cognition, 17*(8), 1271-1286.

Bandura, A., & Walters, R. H. (1977). Social learning theory.

Baron-Cohen, S. (1996). Reading the mind in the face: A cross-cultural and developmental study. *Visual Cognition, 3*(1), 39-60.

Baron-Cohen, S. (2003). *Mind reading [: the interactive guide to emotions*: Jessica Kingsley Publishers.

Baron-Cohen, S., Cox, A., Baird, G., Swettenham, J., Nightingale, N., Morgan, K., . . . Charman, T. (1996). Psychological markers in the detection of autism in infancy in a large population. *The British Journal of Psychiatry, 168*(2), 158-163.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*(1), 37-46.

Baron-Cohen, S., Ring, H., Moriarty, J., Schmitz, B., Costa, D., & Ell, P. (1994). Recognition of mental state terms. Clinical findings in children with autism and a functional neuroimaging study of normal adults. *The British Journal of Psychiatry, 165*(5), 640-649.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of child psychology and psychiatry, 42*(2), 241-251.

Baron-Cohen, S., Wheelwright, S., Jolliffe, & Therese. (1997). Is there a" language of the eyes"? Evidence from normal adults, and adults with autism or Asperger syndrome. *Visual Cognition, 4*(3), 311-331.

Barrett, L. F., & Gendron, M. (2016). The importance of context: Three corrections to Cordaro, Keltner, Tshering, Wangchuk, and Flynn (2016).

Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science, 20*(5), 286-290.

Baumeister, R. F., & Bratslavsky, E. (1999). Passion, intimacy, and time: Passionate love as a function of change in intimacy. *Personality and social psychology review, 3*(1), 49-67.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of general psychology, 5*(4), 323.

Bavelas, J. B., & Chovil, N. (1997). Faces in dialogue.

Bavelas, J. B., & Chovil, N. (2000). Visible acts of meaning an integrated message model of language in face-to-face dialogue. *Journal of Language and social Psychology, 19*(2), 163-194.

Bavelas, J. B., Coates, L., & Johnson, T. (2000). Listeners as co-narrators. *Journal of personality and social psychology, 79*(6), 941.

Becker-Asano, C., & Wachsmuth, I. (2010). Affective computing with primary and secondary emotions in a virtual human. *Autonomous Agents and Multi-Agent Systems, 20*(1), 32-49.

Bell, C. (1824). *Essays on the Anatomy and Philosophy of Expression*: J. Murray.

Berry, D. S., & Pennebaker, J. W. (1993). Nonverbal and verbal emotional expression and health. *Psychotherapy and psychosomatics, 59*(1), 11-19.

Birdwhistell, R. L. (1970/2010). *Kinesics and context: Essays on body motion communication*: University of Pennsylvania press.

Bond, M. H., Wan, K.-C., Leung, K., & Giacalone, R. A. (1985). How are responses to verbal insult related to cultural collectivism and power distance? *Journal of Cross-Cultural Psychology, 16*(1), 111-127.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*: OUP Oxford.

Boukricha, H., Wachsmuth, I., Hofstätter, A., & Grammer, K. (2009). *Pleasure-arousal-dominance driven facial expression simulation.* Paper presented at the Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on.

Bradbury, J. W., & Vehrencamp, S. L. (1998). Principles of animal communication.

Breazeal, C. (2003). Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies, 59*(1), 119-155.

Breazeal, C. L. (2004). *Designing sociable robots*: MIT press.

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental psychology, 18*(6), 906.

Broadbent, E., Kuo, I. H., Lee, Y. I., Rabindran, J., Kerse, N., Stafford, R., & MacDonald, B. A. (2010). Attitudes and reactions to a healthcare robot. *Telemedicine and e-Health, 16*(5), 608-613.

Brown, J. R., Donelan‐McCall, N., & Dunn, J. (1996). Why talk about mental states? The significance of children's conversations with friends, siblings, and mothers. *Child development, 67*(3), 836-849.

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

Bruhn, J. G. (1978). The doctor's touch: tactile communication in the doctor-patient relationship. *Southern medical journal, 71*(12), 1469-1473.

Bruner, J. (1985). Child's talk: Learning to use language. *Child Language Teaching and Therapy, 1*(1), 111-114.

Brunner, L. J. (1979). Smiles can be back channels. *Journal of personality and social psychology, 37*(5), 728.

Bugental, D. E., Love, L. R., Kaswan, J. W., & April, C. (1971). Verbal-nonverbal conflict in parental messages to normal and disturbed children. *Journal of abnormal psychology, 77*(1), 6.

Bull, P., & Frederikson, L. (2001). Nonverbal communication. *The Psychologist, 14*, 644-647.

Bulwer, J. (1649). *Pathomyotomia*.

Caldara, R., Schyns, P., Mayer, E., Smith, M. L., Gosselin, F., & Rossion, B. (2005). Does prosopagnosia take the eyes out of face representations? Evidence for a defect in representing diagnostic facial information following brain damage. *Journal of Cognitive Neuroscience, 17*(10), 1652-1666.

Calder, A. J., Rowland, D., Young, A. W., Nimmo-Smith, I., Keane, J., & Perrett, D. I. (2000). Caricaturing facial expressions. *Cognition, 76*(2), 105-146.

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience, 6*(8), 641-651.

Cang, L. (2012). Grammar of Mental State Terms in Modern Chinese. *Journal of Chinese Language, 2*. doi:1672-8610(2012)2-0018-02

Cangelosi, A., & Parisi, D. (2012). *Simulating the evolution of language*: Springer Science & Business Media.

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

Carton, J. S., Kessler, E. A., & Pape, C. L. (1999). Nonverbal decoding skills and relationship well-being in adults. *Journal of Nonverbal Behavior, 23*(1), 91-100.

Cassell, J. (2000). Embodied conversational agents: MIT press.

Cassell, J., Bickmore, T., Campbell, L., Vilhjalmsson, H., & Yan, H. (2001). More than just a pretty face: conversational protocols and the affordances of embodiment. *Knowledge-based systems, 14*(1), 55-64.

Cassell, J., Sullivan, J., Prevost, S., & Churchill, E. (2000). *Embodied conversational agents*: MIT press.

Celani, G., Battacchi, M. W., & Arcidiacono, L. (1999). The understanding of the emotional meaning of facial expressions in people with autism. *Journal of autism and developmental disorders, 29*(1), 57-66.

Cellan-Jones, R. (2014). Stephen Hawking warns artificial intelligence could end mankind. *BBC News, 2*.

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

Chen, C., Crivelli, C., Garrod, O., Fernandez-Dols, J.-M., Schyns, P., & Jack, R. (2016). Facial Expressions of Pain and Pleasure are Highly Distinct. *Journal of Vision, 16*(12).

Chen, J. (2009). A Review of Chinese Mental State Terms. *Journal of Modern Chinese Language, 2009*(12), 11-12.

Chen, L. (2008). *Non-cluster Analysis of Chinese Mental State Terms.* East China Normal University.

Chen, P. (1999). *Modern Chinese: history and sociolinguistics*: Cambridge University Press.

Chisholm, R. M. (1966). Theory of knowledge.

Chomsky, N. (2014). *Aspects of the Theory of Syntax* (Vol. 11): MIT press.

Chovil, N. (1991). Discourse-oriented facial displays in conversation. *Research on Language & Social Interaction, 25*(1-4), 163-194.

Clark, H. H. (1996). *Using language* (Vol. 1996): Cambridge university press Cambridge.

Clark, H. H., & Tree, J. E. F. (2002). Using uh and um in spontaneous speaking. *Cognition, 84*(1), 73-111.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*(1), 1-39.

Clyde, D. J. (1963). *Manual for the Clyde mood scale*: Biometric Laboratory, University of Miami.

Cohen, J. (1988). Statistical power analysis for the behavior science. *Lawrance Eribaum Association*.

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science, 1*(3), 98-101.

Craig, K. D., Prkachin, K. M., & Grunau, R. V. (1992). The facial expression of pain.

Croft, W. (1993). Case marking and the semantics of mental verbs *Semantics and the Lexicon* (pp. 55-72): Springer.

Cunningham, D. W., Breidt, M., Kleiner, M., Wallraven, C., & Bülthoff, H. H. (2003). *How believable are real faces? Towards a perceptual basis for conversational animation.* Paper presented at the Computer Animation and Social Agents, 2003. 16th International Conference on.

Cunningham, D. W., Kleiner, M., Bülthoff, H. H., & Wallraven, C. (2004). *The components of conversational facial expressions.* Paper presented at the Proceedings of the 1st Symposium on Applied perception in graphics and visualization.

Cunningham, D. W., & Wallraven, C. (2009). Dynamic information for the recognition of conversational expressions. *Journal of Vision, 9*(13), 7.

d'Andrade, R. (1987). A folk model of the mind.

Dario, P., Guglielmelli, E., Laschi, C., & Teti, G. (1999). MOVAID: a personal robot in everyday life of disabled and elderly people. *Technology and Disability, 10*(2), 77-93.

Darwin, C. (1872/1965). *The expression of the emotions in man and animals* (Vol. 526): University of Chicago press.

Darwin, C. (1872/1998). *The expression of the emotions in man and animals* (3rd ed.): Oxford University Press, USA.

Davis, M., & Whalen, P. J. (2001). The amygdala: vigilance and emotion. *Molecular psychiatry, 6*(1), 13-34.

De Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience, 7*(3), 242-249.

De Leersnyder, J., Mesquita, B., & Kim, H. S. (2011). Where do my emotions belong? A study of immigrants' emotional acculturation. *Personality and Social Psychology Bulletin*, 0146167211399103.

De Rosis, F., Pelachaud, C., Poggi, I., Carofiglio, V., & De Carolis, B. (2003). From Greta's mind to her face: modelling the dynamics of affective states in a conversational embodied agent. *International Journal of Human-Computer Studies, 59*(1), 81-118.

De Ruiter, J. P., Mitterer, H., & Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *language*, 515-535.

Delis, I., Chen, C., Jack, R. E., Garrod, O. G., Panzeri, S., & Schyns, P. G. (2016). Space-by-time manifold representation of dynamic facial expressions for emotion categorization. *Journal of Vision, 16*(8), 14-14.

Descartes, R. (1989). *The passions of the soul*: Hackett Publishing.

Descartes, R., & De Spinoza, B. (1619/1961). *Rules for the Direction of the Mind* (Vol. 1): Bobbs-Merrill Indianapolis.

Diener, E. (1994). Assessing subjective well-being: Progress and opportunities. *Social indicators research, 31*(2), 103-157.

Diener, E., & Chan, M. Y. (2011). Happy people live longer: Subjective well-being contributes to health and longevity. *Applied Psychology: Health and Well-Being, 3*(1), 1-43.

Ding, W. (2008). *A Study on Mental State Terms in Modern Chinese.* Shanghai Normal University.

Dissanayake, E. (1979). An ethological view of ritual and art in human evolutionary history. *Leonardo*, 27-31.

Dittmann, A. T., & Llewellyn, L. G. (1968). Relationship between vocalizations and head nods as listener responses. *Journal of personality and social psychology, 9*(1), 79.

Dong, J.-K., Jin, T.-H., Cho, H.-W., & Oh, S.-C. (1999). The esthetics of the smile: a review of some recent studies. *International Journal of Prosthodontics, 12*(1).

Dowling, P. (2008). *International human resource management: Managing people in a multinational context*: Cengage Learning.

Du, S., Tao, Y., & Martinez, A. M. (2014). Compound facial expressions of emotion. *Proceedings of the National Academy of Sciences, 111*(15), E1454-E1462.

Duchenne, G.-B. (1862/1990). *The mechanism of human facial expression*: Cambridge university press.

Dukas, R. (1998). *Cognitive ecology: the evolutionary ecology of information processing and decision making*: University of Chicago Press.

Duncan, S. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of personality and social psychology, 23*(2), 283.

Duncan, S., & Fiske, D. W. (2015). *Face-to-face interaction: Research, methods, and theory* (Vol. 3): Routledge.

Eagly, A. H. (1995). The science and politics of comparing women and men. *American Psychologist, 50*(3), 145.

Eibl-Eibesfeldt, I. (1979). Human ethology: Concepts and implications for the sciences of man. *Behavioral and brain sciences, 2*, 1-57.

Eibl-Eibesfeldt, I. (1989). Human ethology. *New York*.

Eimer, M., Holmes, A., & McGlone, F. P. (2003). The role of spatial attention in the processing of facial expression: an ERP study of rapid brain responses to six basic emotions. *Cognitive, Affective, & Behavioral Neuroscience, 3*(2), 97-110.

Eisler, F. G. (1968). *Psycholinguistics: Experiments in spontaneous speech*: Academic Press.

Ekman, P. (1971). *Universals and cultural differences in facial expressions of emotion.* Paper presented at the Nebraska symposium on motivation.

Ekman, P. (1972). *Universal and cultural differences in facial expression of emotion.* Paper presented at the Nebraska symposium on motivation.

Ekman, P. (1979). About brows: Emotional and conversational signals. *Human ethology*, 169-202.

Ekman, P. (1984). Expression and the nature of emotion. *Approaches to emotion, 3*, 19-344.

Ekman, P. (1989). The argument and evidence about universals in facial expres-sions. *Handbook of social psychophysiology*, 143-164.

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

Ekman, P. (1993). Facial expression and emotion. *American Psychologist, 48*(4), 384.

Ekman, P. (2004). Emotional and conversational nonverbal signals *Language, knowledge, and representation* (pp. 39-50): Springer.

Ekman, P. (2006). *Darwin and facial expression: A century of research in review*: Ishk.

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

Ekman, P., & Friesen, W. V. (1977). Facial action coding system.

Ekman, P., & Friesen, W. V. (1978a). *Facial Action Coding System: Investigatoris Guide*: Consulting Psychologists Press.

Ekman, P., & Friesen, W. V. (1978b). *Manual for the facial action coding system*: Consulting Psychologists Press.

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

Ekman, P., Friesen, W. V., & Tomkins, S. S. (1971). Facial affect scoring technique: A first validity study. *Semiotica, 3*(1), 37-58.

Ekman, P., & Rosenberg, E. L. (1997). *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*: Oxford University Press.

Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science, 164*(3875), 86-88.

el Kaliouby, R. (2005). Mind-reading machines: automated inference of complex mental states. *Computer Laboratory, University of Cambridge*.

El Kaliouby, R., & Robinson, P. (2005). Real-time inference of complex mental states from facial expressions and head gestures *Real-time vision for human-computer interaction* (pp. 181-200): Springer.

Elfenbein, H. A. (2013). Nonverbal dialects and accents in facial expressions of emotion. *Emotion Review, 5*(1), 90-96.

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

Elfenbein, H. A., Beaupré, M., Lévesque, M., & Hess, U. (2007). Toward a dialect theory: cultural differences in the expression and recognition of posed facial expressions. *Emotion, 7*(1), 131.

Elfenbein, H. A., Mandal, M. K., Ambady, N., Harizuka, S., & Kumar, S. (2002). Cross-cultural patterns in emotion recognition: highlighting design and analytical techniques. *Emotion, 2*(1), 75.

Ellsworth, P. C. (2003). Confusion, concentration, and other emotions of interest: Commentary on Rozin and Cohen (2003).

Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition, 44*(3), 227-240.

Faigin, G. (2012). *The artist's complete guide to facial expression*: Watson-Guptill.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods, 39*(2), 175-191.

Feng, L. (2014). A Syntactic Analysis of Mental State Terms in Modern Chinese. *Chinese Cultural and educational materials*(29), 117-119.

Fernández-Dols, J.-M., Carrera, P., & Crivelli, C. (2011). Facial behavior while experiencing sexual excitement. *Journal of Nonverbal Behavior, 35*(1), 63-71.

Fink, B., Grammer, K., & Matts, P. J. (2006). Visible skin color distribution plays a role in the perception of age, attractiveness, and health in female faces. *Evolution and human behavior, 27*(6), 433-442.

Fink, B., & Matts, P. (2008). The effects of skin colour distribution and topography cues on the perception of female facial age and health. *Journal of the European Academy of Dermatology and Venereology, 22*(4), 493-498.

Foster, M. E., Gaschler, A., Giuliani, M., Isard, A., Pateraki, M., & Petrick, R. P. (2012). *Two people walk into a bar: Dynamic multi-party social interaction with a robot agent.* Paper presented at the 14th ACM international conference on Multimodal interaction (ICMI 2012).

Fridlund, A. J. (2014). *Human facial expression: An evolutionary view*: Academic Press.

Frijda, N. H. (2002). What is pain facial expression for? *Behavioral and brain sciences, 25*(04), 460-460.

Frijda, N. H. (2007). *The laws of emotion*: Lawrence Erlbaum Associates Publishers.

Furrow, D., Moore, C., Davidge, J., & Chiasson, L. (1992). Mental terms in mothers' and children's speech: Similarities and relationships. *Journal of child Language, 19*(03), 617-631.

Geen, T. R., & Tassinary, L. G. (2002). The Mechanization of Emotional Expression in John Bulwer's" Pathomyotomia"(1649). *The American journal of psychology*, 275-299.

Geisler, W. S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological review, 96*(2), 267.

Gibson, E. J. (1969). Principles of perceptual learning and development.

Gifford, R., Ng, C. F., & Wilkinson, M. (1985). Nonverbal cues in the employment interview: Links between applicant qualities and interviewer judgments. *Journal of Applied Psychology, 70*(4), 729.

Gill, D., Garrod, O. G., Jack, R. E., & Schyns, P. G. (2014). Facial movements strategically camouflage involuntary social signals of face morphology. *Psychological Science, 25*(5), 1079-1086.

Gockley, R., Bruce, A., Forlizzi, J., Michalowski, M., Mundell, A., Rosenthal, S., . . . Schultz, A. C. (2005). *Designing robots for long-term social interaction.* Paper presented at the Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on.

Golan, O., Baron-Cohen, S., & Golan, Y. (2008). The 'reading the mind in films' task [child version]: Complex emotion and mental state recognition in children with and without autism spectrum conditions. *Journal of autism and developmental disorders, 38*(8), 1534-1541.

Golan, O., Baron-Cohen, S., & Hill, J. (2006). The Cambridge mindreading (CAM) face-voice battery: Testing complex emotion recognition in adults with and without Asperger syndrome. *Journal of autism and developmental disorders, 36*(2), 169-183.

Goldstein, A. G. (1983). Behavioral scientists' fascination with faces. *Journal of Nonverbal Behavior, 7*(4), 223-255.

Gosselin, F., & Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision research, 41*(17), 2261-2271.

Goubert, L., Vlaeyen, J. W., Crombez, G., & Craig, K. D. (2011). Learning about pain from others: an observational learning account. *The Journal of Pain, 12*(2), 167-174.

Granhag, P. A., & Strömwall, L. A. (2002). Repeated interrogations: verbal and non‐verbal cues to deception. *Applied Cognitive Psychology, 16*(3), 243-257.

Gratch, J., DeVault, D., Lucas, G. M., & Marsella, S. (2015). *Negotiation as a challenge problem for virtual humans.* Paper presented at the International Conference on Intelligent Virtual Agents.

Gratch, J., Rickel, J., André, E., Cassell, J., Petajan, E., & Badler, N. (2002). Creating interactive virtual humans: Some assembly required. *IEEE Intelligent systems, 17*(4), 54-63.

Gray, J. A. (1987). *The psychology of fear and stress* (Vol. 5): CUP Archive.

Green, D., & Swets, J. (1966). Signal detection theory and psychophysics. . *New York: Wiley*.

Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition & emotion, 9*(1), 87-108.

Grunau, R. V., & Craig, K. D. (1987). Pain expression in neonates: facial action and cry. *Pain, 28*(3), 395-410.

Guilford, T., & Dawkins, M. S. (1991). Receiver psychology and the evolution of animal signals. *Animal Behaviour, 42*(1), 1-14.

Haldane, J. B. S. (1927). *A mathematical theory of natural and artificial selection, part V: selection and mutation.* Paper presented at the Mathematical Proceedings of the Cambridge Philosophical Society.

Hall, E. T. (1966). The hidden dimension.

Hall, J. A., & Schmid Mast, M. (2007). Sources of accuracy in the empathic accuracy paradigm. *Emotion, 7*(2), 438.

Han, Q. (2006). *Research on Syntactic and Semantic Features of Chinese Mental State Terms.* Huazhong Normal University.

Hao, L. (1999). Cognitive Interpretation of Chinese Mental State Terms. *Journal of Jiamusi University: Social Science*(5), 40-42.

Hao, W. (2011). *Semantic and Syntactic Features and Misuse of Chinese Mental State Terms.* Guangxi Normal University.

Harness Goodwin, M., & Goodwin, C. (1986). Gesture and coparticipation in the activity of searching for a word. *Semiotica, 62*(1-2), 51-76.

Hass, H. (1970). *Human Animal*: GP Putnam's Sons.

Hasson, O. (2000). Knowledge, information, biases and signal assemblages. *Animal signals. Signalling and signal design in animal communication*, 445-463.

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.

Heeger, D. (1998). Signal detection theory: California.

Hegel, F., Lohse, M., & Wrede, B. (2009). *Effects of visual appearance on the attribution of applications in social robotics.* Paper presented at the RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication.

Herman, L. M., Richards, D. G., & Wolz, J. P. (1984). Comprehension of sentences by bottlenosed dolphins. *Cognition, 16*(2), 129-219.

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

Hiatt, L. M., Harrison, A. M., & Trafton, J. G. (2011). *Accommodating human variability in human-robot teams through theory of mind.* Paper presented at the IJCAI Proceedings-International Joint Conference on Artificial Intelligence.

Hicks, C. L., von Baeyer, C. L., Spafford, P. A., van Korlaar, I., & Goodenough, B. (2001). The Faces Pain Scale–Revised: toward a common metric in pediatric pain measurement. *Pain, 93*(2), 173-183.

Hjortsjö, C.-H. (1969). *Man's face and mimic language*: Studen litteratur.

Hobson, R. P., Ouston, J., & Lee, A. (1988). What's in a face? The case of autism. *British journal of psychology, 79*(4), 441-453.

Hofstede, G. (1984). Cultural dimensions in management and planning. *Asia Pacific journal of management, 1*(2), 81-99.

Hofstede, G. H., & Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*: Sage.

Hughes, S. M., & Nicholson, S. E. (2008). Sex differences in the assessment of pain versus sexual pleasure facial expressions. *Journal of Social, Evolutionary, and Cultural Psychology, 2*(4), 289.

Hyvärinen, A., Karhunen, J., & Oja, E. (2004). *Independent component analysis* (Vol. 46): John Wiley & Sons.

Ibrahimagić-Šeper, L., Čelebić, A., Petričević, N., & Selimović, E. (2006). Anthropometric differences between males and females in face dimensions and dimensions of central maxillary incisors. *Med Glas, 3*, 58-62.

Izard, C. E. (1969). The emotions and emotion constructs in personality and culture research. *Handbook of modern personality theory*, 496-510.

Izard, C. E. (1971). The face of emotion.

Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion-cognition relations.

Izard, C. E. (1994). Innate and universal facial expressions: evidence from developmental and cross-cultural research.

Izard, C. E. (2013). *Human emotions*: Springer Science & Business Media.

Jack, R. E. (2013). Culture and facial expressions of emotion. *Visual Cognition, 21*(9-10), 1248-1286.

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.

Jack, R. E., Caldara, R., & Schyns, P. G. (2012). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology: General, 141*(1), 19.

Jack, R. E., Garrod, O. G., & Schyns, P. G. (2014). Dynamic Facial Expressions of Emotion Transmit an Evolving Hierarchy of Signals over Time. *Current Biology*.

Jack, R. E., Garrod, O. G., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences, 109*(19), 7241-7244.

Jack, R. E., & Schyns, P. G. (2017). Toward a Social Psychophysics of Face Communication. *Annual review of psychology, 68*, 269-297.

Jack, R. E., Sun, W., Delis, I., Garrod, O. G., & Schyns, P. G. (2016). Four not six: Revealing culturally common facial expressions of emotion.

James, W. (1890/2013). *The principles of psychology*: Read Books Ltd.

Jaworska, K., van Rijsbergen, N. J., McNair, S. W., Delis, I., Garrod, O. G., Jack, R. E., . . . Schyns, P. G. (2014). Mental representations of emotional facial expressions are more complex rather than less accurate in older observers. *Journal of Vision, 14*(10), 1385-1385.

Jenssen, T. A. (1979). Display modifiers of Anolis opalinus (Lacertilia: Iguanidae). *Herpetologica*, 21-30.

Jia, J., Wu, Z., Zhang, S., Meng, H. M., & Cai, L. (2014). Head and facial gestures synthesis using PAD model for an expressive talking avatar. *Multimedia tools and applications, 73*(1), 439-461.

Jiang, S. (2004). Research on Chinese Mental State Terms. *Journal of Huangshan University*(5), 87-88.

Zhang, J., & Lu, A. (2007). A Study on Structure and Classification of Chinese Mental State Terms. *Journal of South China Normal University: Social Science*(1), 117-123.

Jingyu, Z. (2004). Mental State Terms and Their Typical Structures in Sentences. *Journal of Sichuan International Studies University, 20*(5), 97-101.

Jinjin, H. (2013). Re-classification and Semantic Analysis of Chinese Mental State Terms. *Journal of Jiangxi University of Science and Technology, 8*(1), 79-81.

Johnson, C. N., & Wellman, H. M. (1980). Children's developing understanding of mental verbs: Remember, know, and guess. *Child development*, 1095-1102.

Jolliffe, I. (2002). *Principal component analysis*: Wiley Online Library.

Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2004). When facial attractiveness is only skin deep. *Perception, 33*(5), 569-576.

Kamei, K., Shinozawa, K., Ikeda, T., Utsumi, A., Miyashita, T., & Hagita, N. (2010). *Recommendation from robots in a real-world retail shop.* Paper presented at the International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business horizons, 53*(1), 59-68.

Kaulard, K., Cunningham, D. W., Bülthoff, H. H., & Wallraven, C. (2012). The MPI facial expression database—a validated database of emotional and conversational facial expressions. *PloS one, 7*(3), e32321.

Keltner, D., Ekman, P., Gonzaga, G. C., & Beer, J. (2003). Facial expression of emotion.

Kendon, A. (2004). *Gesture: Visible action as utterance*: Cambridge University Press.

Khullar, S. C., & Badler, N. I. (2001). Where to look? Automating attending behaviors of virtual human characters. *Autonomous Agents and Multi-Agent Systems, 4*(1-2), 9-23.

Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *Neuroimage, 18*(1), 156-168.

Kim, K.-h. (1999). Phrasal unit boundaries and organization of turns and sequences in Korean conversation. *Human Studies, 22*(2-4), 425-446.

Kinloch, P., & Metge, J. (2014). *Talking past each other: problems of cross cultural communication*: Victoria University Press.

Kinsey, A. C., Pomeroy, W. B., & Martin, C. E. (1948). Sexual behavior in the human male.

Knapp, M. L., Hall, J. A., & Horgan, T. G. (2013). *Nonverbal communication in human interaction*: Cengage Learning.

Kohler, C. G., Turner, T. H., Bilker, W. B., Brensinger, C. M., Siegel, S. J., Kanes, S. J., . . . Gur, R. C. (2003). Facial emotion recognition in schizophrenia: intensity effects and error pattern. *American Journal of Psychiatry, 160*(10), 1768-1774.

Krahmer, E., Ruttkay, Z., Swerts, M., & Wesselink, W. (2002). *Pitch, eyebrows and the perception of focus.* Paper presented at the Speech Prosody 2002, International Conference.

Kubota, N., Nojima, Y., Baba, N., Kojima, F., & Fukuda, T. (2000). *Evolving pet robot with emotional model.* Paper presented at the Evolutionary Computation, 2000. Proceedings of the 2000 Congress on.

Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*: Penguin.

LaBarre, W. (1947). The cultural basis of emotions and gestures. *Journal of Personality, 16*(1), 49-68.

Lakin, J. L. (2006). Automatic cognitive processes and nonverbal communication. *The Sage handbook of nonverbal communication*, 59-77.

Lan, G. (2014). A Summary Review of Research on the Nature of Mental States in Modern Chinese. *Journal of Modern Chinese Language*(1), 19-22.

Langland, R. M., & Panicucci, C. L. (1982). Effects of touch on communication with elderly confused clients. *Journal of Gerontological Nursing, 8*(3), 152-155.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and emotion, 24*(8), 1377-1388.

Lawrence, D. H. (1949). Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. *Journal of experimental psychology, 39*(6), 770.

Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature, 401*(6755), 788-791.

Lee, D. H., Mirza, R., Flanagan, J. G., & Anderson, A. K. (2014). Optical origins of opposing facial expression actions. *Psychological Science, 25*(3), 745-752.

Lee, L., Harkness, K. L., Sabbagh, M. A., & Jacobson, J. A. (2005). Mental state decoding abilities in clinical depression. *Journal of affective disorders, 86*(2), 247-258.

Li, L., Liu, X., & Herr, K. (2007). Postoperative pain intensity assessment: a comparison of four scales in Chinese adults. *Pain Medicine, 8*(3), 223-234.

Li, T., & Zhou, Y. (2015). A Study on the Semantic Collocation of Mental State Terms. *Journal of Yunnan Normal University: Research and Teaching Chinese as Second Language, 13*(2), 10-16.

Li, X. (2013). *Matching between Chinese Mental State Terms and Modal Particle at the End of the Sentence.* Huazhong Normal University.

Lim, N. (2016). Cultural differences in emotion: differences in emotional arousal level between the East and the West. *Integrative Medicine Research, 5*(2), 105-109.

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *The Journal of Neuroscience, 7*(11), 3416-3468.

Lynn, S. K., & Barrett, L. F. (2014). “Utilizing” signal detection theory. *Psychological Science, 25*(9), 1663-1673.

Lyons, M. J., Akamatsu, S., Kamachi, M., Gyoba, J., & Budynek, J. (1998). The Japanese female facial expression (JAFFE) database.

Ma, J. (1898). Mashi Wentong (Ma's Grammar): Beijing: The Commercial Press.

Magga, O. H. (2006). Diversity in Saami terminology for reindeer, snow, and ice. *International Social Science Journal, 58*(187), 25-34.

Mahmoud, M., & Robinson, P. (2011). Interpreting hand-over-face gestures *Affective Computing and Intelligent Interaction* (pp. 248-255): Springer.

Marcus, S. (1982). Smile and Suspicion. *Multimedial Communication, 8*, 77.

Marieb, E. N., & Hoehn, K. (2007). *Human anatomy & physiology*: Pearson Education.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological review, 98*(2), 224.

Marsh, A. A., Elfenbein, H. A., & Ambady, N. (2003). Nonverbal “accents” cultural differences in facial expressions of emotion. *Psychological Science, 14*(4), 373-376.

Maslow, A. H., Frager, R., Fadiman, J., McReynolds, C., & Cox, R. (1970). *Motivation and personality* (Vol. 2): Harper & Row New York.

Masters, W. H., & Masters, V. J. (1986). *Human sexual response*: Bantam Books.

Matsumoto, D. (1989). Cultural influences on the perception of emotion. *Journal of Cross-Cultural Psychology, 20*(1), 92-105.

Matsumoto, D. (1990). Cultural similarities and differences in display rules. *Motivation and Emotion, 14*(3), 195-214.

Matsumoto, D. (1992). American-Japanese cultural differences in the recognition of universal facial expressions. *Journal of Cross-Cultural Psychology, 23*(1), 72-84.

Matsumoto, D. (2006). Culture and nonverbal behavior. *Handbook of nonverbal communication*, 219-235.

Matsumoto, D., Olide, A., Schug, J., Willingham, B., & Callan, M. (2009). Cross-cultural judgments of spontaneous facial expressions of emotion. *Journal of Nonverbal Behavior, 33*(4), 213-238.

Matsumoto, D., & Willingham, B. (2006). The thrill of victory and the agony of defeat: spontaneous expressions of medal winners of the 2004 Athens Olympic Games. *Journal of personality and social psychology, 91*(3), 568.

Mátyássy, A., Kelemen, O., Sárközi, Z., Janka, Z., & Kéri, S. (2006). Recognition of complex mental states in patients with alcoholism after long-term abstinence. *Alcohol and Alcoholism, 41*(5), 512-514.

McNair, D. M., & Lorr, M. (1964). An analysis of mood in neurotics. *The journal of abnormal and social psychology, 69*(6), 620.

McNeill, D. (2008). *Gesture and thought*: University of Chicago Press.

Mehrabian, A., & Wiener, M. (1967). Decoding of inconsistent communications. *Journal of personality and social psychology, 6*(1), 109.

Melzack, R., & Wall, P. D. (1967). Pain mechanisms: a new theory. *Survey of Anesthesiology, 11*(2), 89-90.

Messinger, D. S. (2002). Positive and negative: Infant facial expressions and emotions. *Current Directions in Psychological Science, 11*(1), 1-6.

Meston, C. M., Levin, R. J., Sipski, M. L., Hull, E. M., & Heiman, J. R. (2004). Women's orgasm. *Annual review of sex research, 15*(1), 173-257.

Moore, C., Bryant, D., & Furrow, D. (1989). Mental terms and the development of certainty. *Child development*, 167-171.

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine, 19*(2), 98-100.

Morris, D. (1967/2010). *The naked ape: A zoologist's study of the human animal*: Random House.

Morris, J. S., Frith, C. D., Perrett, D. I., & Rowland, D. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature, 383*(6603), 812.

Motley, M. T., & Camden, C. T. (1988). Facial expression of emotion: A comparison of posed expressions versus spontaneous expressions in an interpersonal communication setting. *Western Journal of Communication (includes Communication Reports), 52*(1), 1-22.

Murphy, G. (2004). *The big book of concepts*: MIT press.

Nelson, N. L., & Russell, J. A. (2013). Universality revisited. *Emotion Review, 5*(1), 8-15.

Nesse, R. M. (2005). Natural selection and the regulation of defenses: A signal detection analysis of the smoke detector principle. *Evolution and human behavior, 26*(1), 88-105.

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human brain mapping, 15*(1), 1-25.

Norman, W. T. (1967). 2800 PERSONALITY TRAIT DESCRIPTORS--NORMATIVE OPERATING CHARACTERISTICS FOR A UNIVERSITY POPULATION.

Nowicki, S., & Duke, M. P. (1992). The association of children's nonverbal decoding abilities with their popularity, locus of control, and academic achievement. *The Journal of genetic psychology, 153*(4), 385-393.

Nowlis, V., & Nowlis, H. H. (1956). The description and analysis of mood. *Annals of the New York Academy of Sciences, 65*(4), 345-355.

Nusseck, M., Cunningham, D. W., Wallraven, C., & Bülthoff, H. H. (2008). The contribution of different facial regions to the recognition of conversational expressions. *Journal of Vision, 8*(8), 1.

Nuyts, J., & Pederson, E. (1999). *Language and conceptualization* (Vol. 1): Cambridge University Press.

O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D. M., & Dolan, R. J. (2003). Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia, 41*(2), 147-155.

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: a threat advantage with schematic stimuli. *Journal of personality and social psychology, 80*(3), 381.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, 105*(32), 11087-11092.

Osgood, C. E. (1966). Dimensionality of the semantic space for communication via facial expressions. *Scandinavian journal of Psychology, 7*(1), 1-30.

Paik, Y., & Teagarden, M. B. (1995). Strategic international human resource management approaches in the maquiladora industry: a comparison of Japanese, Korean and US firms. *International Journal of Human Resource Management, 6*(3), 568-587.

Pammi, S., & Schröder, M. (2009). *Annotating meaning of listener vocalizations for speech synthesis.* Paper presented at the 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops.

Paulhus, D. L. (1991). Measurement and control of response bias.

Pelachaud, C., Badler, N. I., & Steedman, M. (1996). Generating facial expressions for speech. *Cognitive Science, 20*(1), 1-46.

Pepperberg, I. M. (1981). Functional vocalizations by an African Grey parrot (Psittacus erithacus). *Zeitschrift für Tierpsychologie, 55*(2), 139-160.

Peterson, R. A. (2000). A meta-analysis of variance accounted for and factor loadings in exploratory factor analysis. *Marketing Letters, 11*(3), 261-275.

Philippot, P. (1993). Inducing and assessing differentiated emotion-feeling states in the laboratory. *Cognition and emotion, 7*(2), 171-193.

Poggi, I., & Pelachaud, C. (2000). Performative facial expressions in animated faces. *Embodied conversational agents*, 155-189.

Porter, S., & Ten Brinke, L. (2008). Reading between the lies identifying concealed and falsified emotions in universal facial expressions. *Psychological Science, 19*(5), 508-514.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and brain sciences, 1*(04), 515-526.

Prkachin, K. M. (1992). The consistency of facial expressions of pain: a comparison across modalities. *Pain, 51*(3), 297-306.

Prkachin, K. M., Berzins, S., & Mercer, S. R. (1994). Encoding and decoding of pain expressions: a judgement study. *Pain, 58*(2), 253-259.

Prkachin, K. M., Solomon, P., Hwang, T., & Mercer, S. R. (2001). Does experience influence judgements of pain behaviour? Evidence from relatives of pain patients and therapists. *Pain Research and Management, 6*(2), 105-112.

Qi, Q. (2007). *A Study on the Internal Differences of Mental State Terms in Modern Chinese.* Anhui Normal University.

Qiao, Y. (2015). Characteristics of Mental State Terms in Noun - Verb Structure. *Journal of Harbin Normal University: Social Science*(2), 79-81.

Reisenzein, R., Studtmann, M., & Horstmann, G. (2013). Coherence between emotion and facial expression: Evidence from laboratory experiments. *Emotion Review, 5*(1), 16-23.

Richoz, A.-R., Jack, R. E., Garrod, O. G., Schyns, P. G., & Caldara, R. (2015). Reconstructing dynamic mental models of facial expressions in prosopagnosia reveals distinct representations for identity and expression. *Cortex, 65*, 50-64.

Rickel, J., Marsella, S., Gratch, J., Hill, R., Traum, D., & Swartout, W. (2002). *Toward a new generation of virtual humans for interactive experiences*. Retrieved from

Riggio, R. E. (2005). Business applications of nonverbal communication. *Applications of Nonverbal Communication. Mahwah, NJ: Lawrence Erlbaum Associates*, 119-138.

Rijsbergen, N., Ince, R., Rousselet, G., Gross, J., & Schyns, P. (2016). Dynamic flow of Face Categorization Task Information in an MEG Network. *Journal of Vision, 16*(12), 1235-1235.

Rosenblatt, A., Greenberg, J., Solomon, S., Pyszczynski, T., & Lyon, D. (1989). Evidence for terror management theory: I. The effects of mortality salience on reactions to those who violate or uphold cultural values. *Journal of personality and social psychology, 57*(4), 681.

Rosmus, C., Johnston, C. C., Chan-Yip, A., & Yang, F. (2000). Pain response in Chinese and non-Chinese Canadian infants: is there a difference? *Social Science & Medicine, 51*(2), 175-184.

Rozin, P., & Cohen, A. B. (2003). High frequency of facial expressions corresponding to confusion, concentration, and worry in an analysis of naturally occurring facial expressions of Americans. *Emotion, 3*(1), 68.

Rozin, P., & Fallon, A. E. (1987). A perspective on disgust. *Psychological review, 94*(1), 23.

Rozin, P., Haidt, J., & McCauley, C. R. (2008). Disgust.

Run, G., Zhang, L., Zhang, X., & Sun, S. (2012). Mechanism of Processing Chinese Mental State Terms in Reading. *Psychology and Behavior Studies, 10*(3), 183-189.

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology, 39*(6), 1161.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies. *Psychological Bulletin, 115*(1), 102.

Russell, J. A. (1997). Reading emotion from and into faces: Resurrecting a dimensional-contextual perspective. *The psychology of facial expression*, 295-320.

Russell, J. A., & Fernández-Dols, J. M. (1997a). *The psychology of facial expression*: Cambridge university press.

Russell, J. A., & Fernández-Dols, J. M. (1997b). What does a facial expression mean. *The psychology of facial expression, 1*, 3.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *language*, 696-735.

Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: an fMRI study. *Cognitive Brain Research, 20*(1), 81-91.

Schegloff, E. A. (1982). Discourse as an interactional achievement: Some uses of 'uh huh'and other things that come between sentences. *Analyzing discourse: Text and talk, 71*, 93.

Schegloff, E. A. (2000). Overlapping talk and the organization of turn-taking for conversation. *Language in society, 29*(01), 1-63.

Scheib, J. E., Gangestad, S. W., & Thornhill, R. (1999). Facial attractiveness, symmetry and cues of good genes. *Proceedings of the Royal Society of London B: Biological Sciences, 266*(1431), 1913-1917.

Scherer, K. R., Mortillaro, M., & Mehu, M. (2013). Understanding the mechanisms underlying the production of facial expression of emotion: A componential perspective. *Emotion Review, 5*(1), 47-53.

Schlosberg, H. (1952). The description of facial expressions in terms of two dimensions. *Journal of experimental psychology, 44*(4), 229.

Schubert, E. F. (2006). Human eye sensitivity and photometric quantities. *Light-Emitting Diodes*, 275-291.

Schultz, D., Izard, C. E., & Ackerman, B. P. (2000). Children’s anger attribution bias: Relations to family environment and social adjustment. *Social Development, 9*(3), 284-301.

Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science, 13*(5), 402-409.

Schyns, P. G., Goldstone, R. L., & Thibaut, J.-P. (1998). The development of features in object concepts. *Behavioral and brain sciences, 21*(01), 1-17.

Scott‐Phillips, T. C. (2008). Defining biological communication. *Journal of evolutionary biology, 21*(2), 387-395.

Scullion, H., Collings, D. G., & Gunnigle, P. (2007). International human resource management in the 21st century: emerging themes and contemporary debates. *Human Resource management journal, 17*(4), 309-319.

Shannon, C. E. (1948/2001). A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review, 5*(1), 3-55.

Shannon, C. E., & Weaver, W. (1949/2015). *The mathematical theory of communication*: University of Illinois press.

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

Shatz, M., Wellman, H. M., & Silber, S. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition, 14*(3), 301-321.

Shweder, R. A., Haidt, J., Horton, R., & Joseph, C. (2008). The cultural psychology of the emotions. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions*. New York: The Guilford press.

Slabbekoorn, H., & den Boer-Visser, A. (2006). Cities change the songs of birds. *Current Biology, 16*(23), 2326-2331.

Smith, F. W., & Schyns, P. G. (2009). Smile through your fear and sadness transmitting and identifying facial expression signals over a range of viewing distances. *Psychological Science, 20*(10), 1202-1208.

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

Song, G., Yin, K., Zhou, Y., & Cheng, X. (2009). A surveillance robot with hopping capabilities for home security. *IEEE Transactions on Consumer Electronics, 55*(4), 2034-2039.

Sorce, J. F., Emde, R. N., Campos, J. J., & Klinnert, M. D. (1985). Maternal emotional signaling: Its effect on the visual cliff behavior of 1-year-olds. *Developmental psychology, 21*(1), 195.

Sowa, J. F. (2014). *Principles of Semantic Networks: Explorations in the representation of knowledge*: Morgan Kaufmann.

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

Sprengelmeyer, R., Young, A., Mahn, K., Schroeder, U., Woitalla, D., Büttner, T., . . . Przuntek, H. (2003). Facial expression recognition in people with medicated and unmedicated Parkinson’s disease. *Neuropsychologia, 41*(8), 1047-1057.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior research methods, instruments, & computers, 31*(1), 137-149.

Stephens, J., & Beattie, G. (1986). On judging the ends of speaker turns in conversation. *Journal of Language and social Psychology, 5*(2), 119-134.

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication, 42*(4), 73-93.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., . . . Yoon, K.-E. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences, 106*(26), 10587-10592.

Sun, L. (2008). Mental State Terms and Their Gramar Features in Chinese Language. *Journal of Hunan Medical University, 10*(3), 125-127.

Susskind, J., Littlewort, G., Bartlett, M., Movellan, J., & Anderson, A. (2007). Human and computer recognition of facial expressions of emotion. *Neuropsychologia, 45*(1), 152-162.

Susskind, J. M., Lee, D. H., Cusi, A., Feiman, R., Grabski, W., & Anderson, A. K. (2008). Expressing fear enhances sensory acquisition. *Nature neuroscience, 11*(7), 843-850.

Symons, D. (1980). The evolution of human sexuality revisited. *Behavioral and brain sciences, 3*(02), 203-214.

Symons, D. K. (2004). Mental state discourse, theory of mind, and the internalization of self–other understanding. *Developmental Review, 24*(2), 159-188.

Tager-Flusberg, H., & Sullivan, K. (1995). Attributing mental states to story characters: A comparison of narratives produced by autistic and mentally retarded individuals. *Applied Psycholinguistics, 16*(03), 241-256.

Tager-Flusberg, H., & Sullivan, K. (2000). A componential view of theory of mind: Evidence from Williams syndrome. *Cognition, 76*(1), 59-90.

Tager-Flusberg, H. (1992). Autistic children's talk about psychological states: Deficits in the early acquisition of a theory of mind. *Child development, 63*(1), 161-172.

Takeuchi, A., & Nagao, K. (1993). *Communicative facial displays as a new conversational modality.* Paper presented at the Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems.

Tanaka, H. (2000). *Turn-taking in Japanese conversation: A study in grammar and interaction* (Vol. 56): John Benjamins Publishing.

Tapscott, D. (1996). *The digital economy: Promise and peril in the age of networked intelligence* (Vol. 1): McGraw-Hill New York.

Taumoepeau, M., & Ruffman, T. (2006). Mother and infant talk about mental states relates to desire language and emotion understanding. *Child development, 77*(2), 465-481.

Tedeschi, J. T., Gaes, G. G., & Rivera, A. N. (1977). Aggression and the use of coercive power. *Journal of Social Issues, 33*(1), 101-125.

Tiedens, L. Z., & Fragale, A. R. (2003). Power moves: complementarity in dominant and submissive nonverbal behavior. *Journal of personality and social psychology, 84*(3), 558.

Tinbergen, N. (1952). "Derived" activities; their causation, biological significance, origin, and emancipation during evolution. *The Quarterly review of biology, 27*(1), 1-32.

Ting-Toomey, S. (1994). *The challenge of facework: Cross-cultural and interpersonal issues*: SUNY Press.

Todorov, A., Baron, S. G., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: a model based approach. *Social cognitive and affective neuroscience, 3*(2), 119-127.

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in cognitive sciences, 12*(12), 455-460.

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

Tomkins, S. S. (1984). Affect theory. *Approaches to emotion, 163*, 195.

Tracy, J. L. (2014). An evolutionary approach to understanding distinct emotions. *Emotion Review, 6*(4), 308-312.

Triandis, H. C. (1995). *Individualism & collectivism*: Westview press.

Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity*: Oxford University Press.

Tsai, J. L., Knutson, B., & Fung, H. H. (2006). Cultural variation in affect valuation. *Journal of personality and social psychology, 90*(2), 288.

Tung, R. L. (1982). Selection and training procedures of US, European, and Japanese multinationals. *California Management Review, 25*(1), 57-71.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind, 59*(236), 433-460.

Van Rijsbergen, N., Jaworska, K., Rousselet, G. A., & Schyns, P. G. (2014). With age comes representational wisdom in social signals. *Current Biology, 24*(23), 2792-2796.

Wagner, H. L., MacDonald, C. J., & Manstead, A. (1986). Communication of individual emotions by spontaneous facial expressions. *Journal of personality and social psychology, 50*(4), 737.

Wall, P. D. (1978). The gate control theory of pain mechanisms. A re-examination and re-statement. *Brain, 101*(1), 1-18.

Wall, P. D. (2000). *Pain: the science of suffering*: Columbia University Press.

Waller, B. M., Cray Jr, J. J., & Burrows, A. M. (2008). Selection for universal facial emotion. *Emotion, 8*(3), 435.

Wallraven, C., Hur, D.-C., & Shin, A. (2015). Across Cultures: A Cognitive and Computational Analysis of Emotional and Conversational Facial Expressions in Germany and Korea *Recent Progress in Brain and Cognitive Engineering* (pp. 97-108): Springer.

Walters, E. T. (1994). Injury-related behavior and neuronal plasticity: an evolutionary perspective on sensitization, hyperalgesia, and analgesia. *International review of neurobiology, 36*, 325-427.

Wang, H. (2002). The Classification of Mental State Terms in Modern Chinese. *Journal of Jinsan Normal College, 19*(4), 62-64.

Wang, H. (2004). Mental State Terms in event-related and other sentences. *Journal of Nanjing Normal University: Social Science*(2), 139-144.

Wang, Y. (2012). A Review of Studies on Chinese Mental State Terms. *Journal of Mudanjiang Normal College: Philosophy and Social Science*(3), 82-84.

Warkentin, M. E., Sayeed, L., & Hightower, R. (1997). Virtual teams versus face-to-face teams: an exploratory study of a Web-based conference system. *Decision Sciences, 28*(4), 975-996.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior research methods, 45*(4), 1191-1207.

Wasti, S. A. (1998). Cultural barriers in the transferability of Japanese and American human resources practices to developing countries: The Turkish case. *International Journal of Human Resource Management, 9*(4), 608-631.

Waters, K. (1987). A muscle model for animation three-dimensional facial expression. *ACM Siggraph Computer Graphics, 21*(4), 17-24.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*(2), 219.

Wen, Y. (2007). Mental State Terms in Mordern Chinese.

Widen, S. C. (2013). Children's interpretation of facial expressions: The long path from valence-based to specific discrete categories. *Emotion Review, 5*(1), 72-77.

Widen, S. C., & Russell, J. A. (2008). Children acquire emotion categories gradually. *Cognitive development, 23*(2), 291-312.

Wierzbicka, A. (1999). *Emotions across languages and cultures: Diversity and universals*: Cambridge University Press.

Williams, A. C. d. C. (2002). Facial expression of pain, empathy, evolution, and social learning. *Behavioral and brain sciences, 25*(04), 475-480.

Wilson, E. O. (2012). *The social conquest of earth*: WW Norton & Company.

Woodzicka, J. A., & LaFrance, M. (2005). Working on a smile: Responding to sexual provocation in the workplace. *Applications of nonverbal communication*, 139-155.

Xu, R., & Wang, W. (2005). Analysis on Mental State Terms in Chinese Language. *Journal of Ningbo University: Humanities, 18*(3), 65-69.

Yacoob, Y., & Davis, L. S. (1996). Recognizing human facial expressions from long image sequences using optical flow. *IEEE Transactions on pattern analysis and machine intelligence, 18*(6), 636-642.

Yang, Y. (1999). Mental State Terms Are Independent From Adverbs. *Journal of Yunnan Normal University: Research and Teaching Chinese as Second Language*(1), 71-74.

Yin, L., Chen, X., Sun, Y., Worm, T., & Reale, M. (2008). *A high-resolution 3D dynamic facial expression database.* Paper presented at the Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference On.

Yngve, V. H. (1970). *On getting a word in edgewise.* Paper presented at the Chicago Linguistics Society, 6th Meeting.

You, Z.-J., Shen, C.-Y., Chang, C.-W., Liu, B.-J., & Chen, G.-D. (2006). *A robot as a teaching assistant in an English class.* Paper presented at the Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06).

Yu, H., Garrod, O. G., & Schyns, P. G. (2012). Perception-driven facial expression synthesis. *Computers & Graphics, 36*(3), 152-162.

Zapf, D. (2002). Emotion work and psychological well-being: A review of the literature and some conceptual considerations. *Human resource management review, 12*(2), 237-268.

Zeng, Z., Pantic, M., Roisman, G., & Huang, T. S. (2009). A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 31*(1), 39-58.

Zhang, C. (2012). *The Influence of Mental State Terms on False Beliefs of Chinese Children.* East China Normal University.

Zhang, H., & Liu, W. (2011). Mental State Terms and Their Function in Sentences in Modern Chinese. *Journal of Dalian University, 32*(5), 83-87.

Zhang, S., Wu, Z., Meng, H., & Cai, L. (2007). Facial expression synthesis using PAD emotional parameters for a Chinese expressive avatar. *Affective Computing and Intelligent Interaction*, 24-35.

Zhou, Y., & Shao, J. (1993). Chinese Mental State Terms and Their Function in Sentences. *Journal of Chinese Language*(3), 32-36.

Zhu, L. (2012). *Research on Mental State Terms in Modern Chinese Based on Valence Theory.* Sichuan International Studies University.

Zorn, T. E., & Violanti, M. T. (1996). Communication abilities and individual achievement in organizations. *Management Communication Quarterly, 10*(2), 139-167.