1

Facial Expressions versus Words: Unlocking Complex Emotional Responses of Residents towards Tourists

ABSTRACT

Tourism as an emotional sphere, researchers' efforts on emotions lag behind the fruitful achievements of psychology, in both methods and theories. Tourism studies on emotion mostly rely on self-reports only, thus limiting the understanding to explicitly expressed emotions. This study aims to compare residents' emotional responses toward tourists expressed implicitly (through facial expressions) and explicitly (through self-reports), and interpret identified discrepancies by exploring the psychological mechanism behind the two expression channels. Using self-developed video vignettes as triggers, Hong Kong residents' facial expressions during watching and self-reported emotions after watching were recorded. Through a comprehensive comparison, desires-derived and stereotypes-elicited emotional responses of residents toward tourists were distinguished. Facial expressions conveyed more desires-derived emotions like happy, sad, and angry, whereas self-reports emphasized stereotypes-elicited emotions, particularly disgust. A dual-process model of emotion formation was proposed to interpret the emotional expressive discrepancies, thereby enhancing the theorisation of tourism studies on emotion.

1. INTRODUCTION

Emotions, as important carriers of communicative meaning (Hetland, Vittersø, Fagermo, Øvervoll, & Dahl, 2016), are critical for tourism and hospitality spheres whereby intense and frequent human interactions take place (Koc & Boz, 2020). Observing a shift from disembodied accounts of to an embodiment focus in tourism, the last two decades have witnessed an 'emotional turn' in this field, accompanied by a proliferation of research on emotion (Cohen & Cohen, 2019). While the important role of emotion in determining tourists' experience and satisfaction has been substantially examined, few studies have concerned residents' emotional responses toward tourists due to a business-focused account of tourism (Prayag, 2020). Given the same pivotal role of residents in cultivating hospitable destination social servicescape, positive host-tourist relationship and destination sustainability, research on residents' emotions is timely warranted (Zheng, Ritchie, Benckendorff, & Bao, 2019). Moreover, emotion research in tourism as a whole, even with a rising trend, is still in its infancy compared to that in psychology, in terms of both methods and theoretical interpretations (Hosany, Martin, & Woodside, 2020).

Emotion research in tourism can be traced back to consumer and marketing research, which is grounded in psychology. However, extant tourism research often ignores the complexity of emotions recognised by psychologists and oversimplifies the measurement of emotions by relying on self-reported methods (Prayag, Hosany, Muskat, & Del Chiappa, 2017). In psychology, emotion is widely acknowledged to encompass subjective feelings, expressive motor behaviours, physiological arousal, cognitive appraisal and behavioural tendencies (Frijda, 1986). In this regard, emotion is a complex state that can be expressed through various channels, including implicit ones like facial expressions, physiological changes and bodily reactions, and explicit ones as verbal or written reports (Scherer & Moors, 2019). Implicitly and explicitly expressed emotions are usually distinguished as automatic and conscious (Baumeister & Bushman, 2020), non-verbal and verbal (Esposito, 2009), or raw and reflective (Nilsen & Kaszniak, 2007); they represent the spontaneous and deliberate responses of individuals underlying unconscious and conscious control, respectively (Scherer & Moors, 2019). In face-to-face encounters, both implicitly and explicitly expressed emotions are critical to human interaction, particularly the implicit part that exerts salient influence on individuals' communication in a subliminal way (Scherer, 2005a). Whereas traditional self-reported methods as explicit measures are solely effective in capturing explicitly expressed emotions, which are subjective to cognitive and social desirability biases (Mauss & Robinson, 2009),

implicit measures, like heart rate and skin conductance detection, have been widely experimented and encouraged in psychology to complement self-reports for unbiased measurements (Calvo & D'Mello, 2010). Yet combination of implicit and explicit measures to record both forms of emotion expressions is still forthcoming in tourism studies (Hadinejad, Moyle, Scott, & Kralj, 2019). The primary objective of this study is, thus, to examine residents' emotional responses toward tourists expressed implicitly and explicitly by using both implicit and explicit measures.

Specifically, facial expression recognition was selected as the comparable measure with self-report because it is one of the most empirically grounded techniques in use. On the basis of Darwin's expression reflexology, the hardwired connections between facial expressions and basic emotions have been theoretically and empirically examined over the past decades (Ekman, 1970, 1993; Keltner, Tracy, Sauter, & Cowen, 2019). Emotion recognition through facial expressions is superior to other implicit measures (e.g., heart rate and skin conductance) in that it can recognise discrete emotions other than arousal levels (Prayag et al., 2017). Moreover, facial expression recognition using cutting-edge computer programmes, such as FaceRearder, allows real-time capture of implicitly expressed emotions (Lewinski, den Uyl, & Butler, 2014). Even though facial expressions can sometimes be consciously masked, the micro-expressions captured by FaceReader within milliseconds of the stimulation are usually spontaneous beyond humans' conscious control (Penn, 2006). Using sophisticated facial expression recognition technique to supplement self-report for full capture of residents' emotions is of significance for both academics and industry in unravelling complex tourism encounters (Hosany et al., 2020).

Some pioneer work in tourism has compared implicit measures with self-reports, however conclusions were made in a descriptive manner without theoretical explanations. For instance, Hadinejad, Moyle, Kralj, and Scott (2019) compared people's physiological and self-reported emotional responses to tourism marketing stimuli, documenting inconsistence in emotional arousal between the two measures but leaving the discrepancy unexplained. Hetland et al. (2016) reported few correlations between facially-expressed and self-reported emotions in responding to tourism advertising films while ignored theoretical interpretations of the uncorrelated part. The identified but disregarded discrepancies in previous studies necessitate further examinations of emotion formation and expression by differentiating their underlying mechanisms. Even in marketing and psychology, these questions remained underexplored. Most efforts integrating implicit and explicit measures have been made to detect the efficacy and agreement between different methods yet lacking theoretical interpretation of the

discrepancies (Scherer & Moors, 2019). Some phycologists have conceptually connected implicit and explicit expressions of emotion to two distinct appraisal systems, one intuitive/automatic and one reflective/conscious (Lazarus, 1991). For instance, Ekman (1992, p. 184) posited automatic and conscious response systems in corresponding to the two different emotional expression forms and postulated 'some coherence, some systematical relationship between these two response systems (expression vs. autonomic changes) during emotion events'. However, to the best our knowledge, no further effort has been made to investigate the how and why, or the exact theoretical rationale by which emotions in the two expressive forms differ and correlate to the two appraisal systems. Hence, the second objective of this study is to fill in this research gap by digging into the discrepancies between facially-expressed and self-reported emotions, and exploring the underlying psychological mechanisms.

To reiterate, this study aims to thoroughly compare emotional responses of residents toward tourists expressed implicitly and explicitly by using technology-assisted (i.e., facial expression recognition) and traditional (i.e., self-report) methods and explore the mechanisms underlying their discrepancies. The main research questions that guide this study are: 'Are there differences between emotional responses measured by facial expression recognition and self-report? If yes, how and why?' By answering these questions, the study can enrich emotion interpretations in tourism and psychology. This study focuses on Hong Kong residents' emotional responses toward Mainland Chinese tourists based on two considerations. First, the widely reported negative attitudes and sentiments of Hong Kong residents toward Mainland Chinese tourists deserve further explorations in real-time emotional encounters (Chen, Hsu, & Li, 2018; Shen, Luo, & Zhao, 2017). Second, Mainland Chinese tourists is one of the fastest-growing tourist markets worldwide and has become the top source market for many international destinations (Cheung & Li, 2019). Digging into Hong Kong residents' emotional responses to Mainland tourists is of significance for many other host communities in achieving favourable host-tourist communications and sustainable tourism development.

2. LITERATURE REVIEW

2.1 Emotion as a Complex Psychological Response

The philosophical foundation of emotion can be traced back to 'passions' discussed by Plato and Aristotle in the Golden Age of Greece. 'Passions' include 'fear, anger and all other similar emotions', as forces competing with Reason (Konstan, 2006). In this regard, emotions, at its

very beginning, were intuitive responses of individuals, opposite of rational thinking (Crivelli & Fridlund, 2019). However, psychological work on emotion in the 20th century discovered the cognitive feature of emotion. For instance, James (1890, cited from Drozdova, 2014) defined emotion as 'feeling of bodily changes' and connected it to conscious judgements (Robinson, 2005). Darwin (1872) postulated emotion as a set of cognitive, physical and psychological responses associated with unique neural structures on an evolutionary basis. Arnold's (1960) cognitive appraisal theory even viewed emotions as bodily and mental reactions depending on cognitive appraisals, thereby depreciating the intuitive part of emotion. Despite these theoretical domains have not reached an agreement on the definition of emotion, they agree that emotion is 'a complex psychological response' that incorporates various components, including 'cognitive component or appraisal; neurophysiological component or bodily symptoms; motivational component or action tendencies; motor expression component or facial and vocal expression; subjective feeling component or emotional experience' (Scherer, 2005b, p. 698). These components, though varying slightly in different theories (Scherer & Moors, 2019), imply that emotion can be expressed implicitly (e.g., through facial expressions and bodily indications) and explicitly (e.g., through subjective deliberations) (Nilsen & Kaszniak, 2007). Implicitly expressed emotions are likely automatic or unconscious responses evoked by stimuli, while explicitly expressed emotions are more of cognitive or conscious responses toward triggers (Baumeister & Bushman, 2020). Both are indispensable in face-to-face communication because people retrieve emotional information from each other in verbal as well as nonverbal languages (Esposito, 2009).

Recognising the complexity of emotion in expressions and the fact that much of an individual's emotional responses is unconscious and automatic (Penn, 2006), research in psychology and marketing has distinguished implicit (or objective/indirect) and explicit (or subjective/direct) measures for emotion detection (De Houwer, 2006). Traditional explicit measures, including verbal or visual self-reported measurements, ask participants to report their felt emotions (Lagast, Gellynck, Schouteten, De Herdt, & De Steur, 2017). While convenient to use, self-reported methods could have cognitive and social desirability bias and thus limit the understanding of emotions to those explicitly expressed (Keltner et al., 2019). The automatic or unconscious emotional responses that can occur in the first microseconds of a stimulation are not perceived by the subjects, thus unlikely to be verbalized (Baumeister & Bushman, 2020). Implicit measurements, such as heart rate and respiration, blood pressure, skin conductance, facial muscle activity, finger temperature, and eye movement variability,

have gained increasing momentum in research for automatic emotion recognition (van Zonneveld, de Sonneville, Van Goozen, & Swaab, 2019). After being widely experimented in psychology and marketing, these implicit measures have been validated as effective and necessary to complement self-reports (Mauss & Robinson, 2009).

Amongst the many implicit measures, facial expression recognition is relatively superior because it can distinguish concrete emotions (Sato, Hyniewska, Minemoto, & Yoshikawa, 2019), whereas other methods such as heart rate and skin conductance can only report the activation level of emotions (Moors, 2009). Moreover, the connections between facial expressions and emotions have a grounded theoretical base: Darwin's (1872) expression reflexology and the corresponding Basic Emotion Theory which is one of the most salient theories in emotion (Crivelli & Fridlund, 2019). According to the theory, six emotions – anger, happiness, surprise, disgust, sadness, and scare – are distinguished as the most basic emotions that associated with specific facial movements by evolution (Ekman, 1970). Tomkins (1963) and his followers, such as Ekman (1970) and Izard (1971), as well as Plutchik (1980) have devoted themselves to providing both empirical evidence and theoretical explanations to the universal facial expressions of basic emotions. Based on the hardwired connections identified between facial expressions and emotions, the past decades have witnessed substantial advancements in facial expression recognition techniques, among which Facial electromyography (fEMG), Facial Action Coding System (FACS) and automated-based face recognition are the three main threads (Wolf, 2015).

Comparatively, the automated techniques (e.g., FaceReader and AFFDEX) based on computer-vision algorithms are favoured because they can overcome the limitations of fEMG which is equipment-intensive and FACS which is labour-intensive (Wolf, 2015). fEMG recognises emotions through the activation of facial muscles and requires electrodes attached to the skin surface. FACS proposed by Ekman and Friesen (1978, cited from Ekman, 1999) links facial muscle movements with discrete emotions, enabling emotion recognition through observation of an individual's facial expression change, which is subjective and time-consuming (Sato et al., 2019). With the aid of computer science, automated technologies are more reliable and efficient because real-time emotions can be reported automatically as long as videos or images of the face of a person exposed to stimuli is recorded (Lewinski et al., 2014). In particular, FaceReader, software marketed by Noldus (www.noldus.com), is widely used since it can differentiate facial changes by culture and age (e.g., East-Asian Model and Baby Face Reader). FaceReader can recognise six basic emotions (i.e., anger, happiness,

surprise, disgust, sadness, and scare) and the neutral state through calculating the facial action units, with higher than 89% accuracy (Lewinski et al., 2014). Even though humans' facial emotions could arguably be masked, the micro-expressions, which appear between 1/25s and 1/5s, can be recognised by FaceReader to unveil the hidden emotional states (Ekman, 1985/2009).

While technological developments have accelerated the triangulation of implicit and explicit measures to capture complex emotional experience, theoretical explanations of the discrepancy between the two measures are still lacking (Lagast et al., 2017; Mauss & Robinson, 2009). Distinct measures of various expressions have been initially hypothesized to be consistent, the majority of research in psychology and marketing has thus focused on the efficacy and agreement between different approaches in measuring emotions (Calvo & D'Mello, 2010; Tran, Siemer, & Joormann, 2011; Walsh, Duncan, Bell, O'Keefe, & Gallagher, 2017). For instance, Walsh et al. (2017) highlighted consistence between participants' emotional responses toward food measured by physiological methods and self-report. Although inconsistences between implicit and explicit measures were also identified, theoretical interpretation of the conflicting results remains unclear (Scherer & Moors, 2019). In accordance with the widely accepted dual-system (i.e., intuitive vs. reflective) of information processing of human mind (Kahneman & Frederick, 2001), some psychologists categorised appraisal systems in attempts to accommodate distinct emotional expressions. For instance, Ekman (1992) proposed two appraisal systems – automatic vs. extended – in emotional response to stimuli. Lazarus (1991, p. 3) also distinguished two modes of appraisal: 'one automatic, unreflective, and unconscious or preconscious, the other deliberate and conscious.' Their categorisations shared a similar idea that there are two different appraisal systems – automatic and conscious – underlying emotion processing and expression.

The automatic appraisal system accounts for the intuitive mind which operates with great speed to enable extraordinarily short intervals between stimuli and emotional responses (Ekman, 1992). Since this system operates on what is biologically given or socially learnt after repeated encounters, the derived emotional responses happen without awareness or efforts (Kahneman & Frederick, 2001). By contrast, the conscious appraisal system relies on reflective judgement that is slow and deliberate, and the generated emotional responses are within individuals' control and self-awareness (Ekman, 1993) so that socially desirable and masking emotional responses could happen (Haidt, 2001). Although the two systems of appraisal imply distinct emotional processing and expressions, how they function to result in emotional

discrepancies in facial expressions and words warrants further explorations in empirical contexts.

2.2 Emotion Research in Tourism

Defined as affective states of a person arising from appraisals of self-relevant interactions with the environment, emotion is deemed essential in understanding human-environment interaction (Prayag, 2020). Tourism, full of multisensory-based interactions between human and environment, is undoubtedly emotional laden (Jordan, Spencer, & Prayag, 2019). However, it is not until the last decade that emotion gained increasing attention in tourism (Nawijn & Biran, 2019). Cohen and Cohen (2019) posited emotion as a forefront topic in future tourism research owing to substantial gaps identified despite the recent proliferation of related publications. First and foremost, the majority of research focused on tourists' emotional responses, yet neglecting the other important stakeholder – hosts – in tourism encounters (Prayag, 2020). Centring on the business aspects of tourism, extant research concentrated on tourists' emotion in defining their experience (Faullant, Matzler, & Mooradian, 2011; Jordan et al., 2019), satisfaction (Prayag et al., 2017) and behavioural intention (Li, 2019; Pestana, Parreira, & Moutinho, 2019). Not until recent years, there appeared some exceptions that concerned residents' emotional responses toward tourism performing arts (Zheng et al., 2019) or employees' emotional recognition ability in service encounters (Boz & Koc, 2019; Lin & Lin, 2017). Though with a host's view, these studies cared about hosts' emotional intelligence or emotional responses toward tourism development rather than tourists. Emotional responses of residents toward tourists remain an underexplored and warranted area of investigation considering the significance of residents in defining host-tourist interaction and destination sustainable development (Moyle, Moyle, Bec, & Scott, 2019; Prayag, 2020).

Furthermore, extant tourism research mainly measures emotions by self-reported methods using questionnaire or interview with semantic differential descriptors (Hadinejad, Moyle, Scott, et al., 2019; Li, Scott, & Walters, 2015). This line of research, as discussed earlier, has overlooked the implicitly expressed emotions. Moreover, when conducting survey research in tourism, the dimensional approach has been favoured over the basic emotion approach (Hosany et al., 2020; Li et al., 2015). Emotion was mostly measured by two dimensions, valence (i.e., positive or negative) and arousal (i.e., activated or non-activated) (Jordan et al., 2019; Li, 2019). While this approach is useful, it hides the various roles of specific emotions in tourism encounters (Prayag et al., 2017). Different emotions entail distinct behavioural tendencies (Ekman, 1999). For instance, although angry and disgust are both negative emotions,

they have distinct behavioural cues, approaching and avoiding, respectively (Fu, 2015; Walsh et al., 2017). Therefore, examining specific emotions in tourism encounters is important for both theoretical development and marketing strategy formulation (Zheng et al., 2019). The limited number of tourism studies that examined specific emotions, unfortunately, only measured one or a few discrete emotions (Nawijn & Biran, 2019). For instance, Faullant et al. (2011) examined joy, fear and guilt as antecedents of satisfaction in mountaineering experience. More efforts with full considerations of basic emotions combining implicit measures in a tourism context are needed (Hosany et al., 2020; Moyle et al., 2019; Pestana et al., 2019).

Several frontier articles in recent years introduced sophisticated technologies such as FaceReader (Hadinejad, Moyle, Scott, et al., 2019), electroencephalography (EEG) or electrodermal activity (EDA) device (Kim & Fesenmaier, 2015; Li, 2019) to measure tourists' automatic emotional responses toward tourism advertisements or during touring experience. Some focused on the usefulness of implicit measures in a tourism context (e.g. González-Rodríguez, Díaz-Fernández, & Gómez, 2020; Li, Walters, Packer, & Scott, 2018b), while others integrated data from psychophysiological measures and self-reports for comparison (Hetland et al., 2016). Nevertheless, the comparisons made are more of descriptive nature, theoretical interpretation of the comparison results and exploration of underlying mechanisms are absent. For instance, Li, Walters, Packer, and Scott (2018a) used both psychophysiological measures (EDA and fEMG) and self-reported measures to explore the influence of ad-evoked emotions on tourism advertising effectiveness. The results demonstrated varied models for different measures but without explanations. Hadinejad, Moyle, Kralj, et al. (2019) compared FaceReader, skin conductance, self-report survey and interviews in understanding tourists' emotional responses to promotional music. Upon identified inconsistence in emotional arousal revealed by physiological and self-report measures, the study highlighted the importance of utilising multiple methods and left the necessity of deepening interpretation of results from different methods to future research.

To bridge these gaps, this study combines facial expression recognition technique (i.e., FaceReader) and self-report to unravel the complexity of emotions in a host-tourist interaction setting by capturing residents' emotional responses expressed implicitly and explicitly. In comparing results from the two measurements, detection of six basic emotions and their intensity levels, as well as the dimensional feature such as valence will be included. Moreover, discrepancies between the results of the two measurements will be interpreted by exploring the psychological mechanisms underpinning the emotion formation and expression process. This

1 means the use of a grounded theoretical induction approach based on discrepancies observed

2 between the two measures in accordance with psychological propositions such as the two

3 appraisal systems.

4

5 3. METHODOLOGY

6 This exploratory study adopts a grounded theory approach by conducting comparison and

7 inductive reasoning to identify and interpret differences in residents' implicit and explicit

8 emotional responses toward tourists (Denzin & Lincoln, 2008). Comparison is a fundamental

9 research method in social sciences. Through comparisons, scholars can find differences,

10 identify gaps and set goals for theory induction (Liu, 2018). It is one of the most important

11 intelligent ways to know the world (Caramani, 2008), thus 'thinking without comparison is

12 unthinkable' (Swanson, 1971, p. 145). According to the levels of analysis, comparative studies

13 can be classified into three types, descriptive, analytical and explanatory. Explanatory

14 comparison is based upon the previous two and goes further to build relationships and causal

15 connections for the differences (Ragin, 1981). It is adopted by this study to better understand

16 residents' emotional responses. Three steps of the explanatory comparison are strictly followed:

17 (1) Describing differences; (2) In-depth analysing the differences, such as categorisation; and

18 (3) Providing theoretical explanations (Smelser, 2013).

19 Specifically, this study aims to compare and interpret the emotional responses of Hong

20 Kong residents toward Mainland Chinese tourists collected from the facial expression

21 recognition technique (i.e., FaceReader) and self-reports. To achieve the research objectives,

22 video vignettes about Mainland Chinese tourists' behaviours were produced to stimulate

23 residents' emotional responses. Facial expressions and self-reported emotions were recorded

24 during and right after the residents viewed the videos, respectively. The real-time facial

25 expressions were analysed by FaceReader to identify automatic emotions, while self-reports

26 measured explicitly expressed emotions that were consciously perceived by residents

27 themselves (Hosany et al., 2020). To facilitate comparison, self-reports focused on

28 investigating the six basic emotions and corresponding intensities, as what the FaceReader did.

29 3.1 Data Collection

30 Data collection was completed in February 2019 following a three-step process. First, the

31 research team produced 10 video vignettes of Mainland Chinese tourists in Hong Kong,

32 following a rigorous development protocol. The scenarios were based on interviews with 20

Hong Kong residents. Among the 57 personal interaction stories with Mainland Chinese tourists reported, 72% were negative (e.g., jumping queues, children urinating or defecating in public) and only 9% were positive. Thus, among the 10 videos produced, seven depict negative/deviant behaviours and three depict positive behaviours (see Table 1 footnote for details). Each video lasts 1-2 minutes. The videos were pre-tested for their realism and authenticity. Second, a panel of Hong Kong permanent residents were invited through purposive and snowball sampling to watch one to three videos in a quiet room. A video camera was placed in front of them, upon their informed consent, to record their facial expressions when watching the videos. The videos were assigned to each participant randomly to avoid stimuli bias (Denzin & Lincoln, 2008). Third, after watching each video, the participants were asked to answer a short survey that contains an adapted Geneva Wheel (Sacharin, Schlegel, & Scherer, 2012) to measure explicit emotions. For each of the emotions experienced, respondents can indicate a value on a five-point Likert-type scale ('1' = 'very weak' and '5' = 'very strong'). In-depth individual interviews were also conducted after watching each video to explore participants' complex self-perceived emotions and interpretations. The major interview questions include 'How do you feel about the scenarios depicted in the video you just watched?' and 'Why do you feel that way?'.

Data saturation was reached after obtaining 29 interviews from 14 participants (see Table 1 for their profiles) who watched 29 videos in total, which were then coded as 29 data cases by combining the participant number and video clip number for anonymity purpose (Denzin & Lincoln, 2008). Each of the 10 videos received at least two views. Comparisons were made among results from the 29 cases. Interviews of the 29 cases lasted an average of 40 minutes. The sample size, though relatively small, is common and reasonable in emotion studies using cutting-edge physiological technologies; such studies can have sample sizes below or around 30 (e.g., Gakhal & Senior, 2008; Hadinejad, Moyle, Kralj, et al., 2019; Kim & Fesenmaier, 2015; Somervuori & Ravaja, 2013). From the demographic characteristics displayed in Table 1, this sample is representative in terms of age and gender, but over-represents highly educated population. All 14 participants had visited the Mainland more than three times and almost half of them have friends or relatives in Mainland China. Although they are all permanent Hong Kong residents sharing the same Chinese ethnicity, their self-identities vary considerably, with seven identified themselves as primarily Hongkongese but also Chinese, five as Hongkongese only, and two as primarily Chinese but also Hongkongese.

Table 1.

Participant Profile

No.	Code: Participant_Video [a]	Age	Gender	Education	Self-identity
1	P1_V3				
2	P1_V8	32	Female	Postgraduate	Hongkongese but also Chinese
3	P1_V10				
4	P2_V1	40	Male	Postgraduate	Hongkongese but also Chinese
5	P2_V7				
6	P3_V3				
7	P3_V7	61	Female	Upper secondary	Hongkongese
8	P3_V9				
9	P4_V4				
10	P4_V5	40	Male	Postgraduate	Chinese but also Hongkongese
11	P4-V8				
12	P5_V5	38	Male	Degree	Hongkongese
13	P5_V6				
14	P6_V4	31	Female	Postgraduate	Hongkongese
15	P7_V5	31	Female	Degree	Hongkongese
16	P7_V8				
17	P8_V4	61	Male	Upper secondary	Hongkongese but also Chinese
18	P8_V9				
19	P9_V2	53	Female	Degree	Hongkongese but also Chinese
20	P9_V3				
21	P10_V7	29	Female	Degree	Hongkongese
22	P11_V7	49	Female	Degree	Chinese but also Hongkongese
23	P12_V9	22	Male	Degree	Hongkongese but also Chinese
24	P12_V10				
25	P13_V2				
26	P13_V6	20	Male	Degree	Hongkongese but also Chinese
27	P13_V10				
28	P14_V1	63	Male	Postgraduate	Hongkongese but also Chinese
29	P14_V2				

Note: [a] V1- a boy urinating in public, V2- politely asking for direction to a shopping mall, V3- trying on cosmetic samples unhygienically in shop, V4- jumping the queue, V5- asking for smoking area patiently, V6- asking for numerous toiletries in hotel, V7- chatting loudly with hotel room door open, V8- drinking and speaking loudly in public transport, V9- a mother providing civil behaviour guidance to her son in public transport, V10- packing suitcase on the street blocking pedestrian walkway.

3.2 Data Analysis

Survey responses were input into SPSS 25.0 for further analysis. Interviews were transcribed verbatim and then coded following a standard procedure of 'open coding – creating categories – abstraction' (Denzin & Lincoln, 2008). Videos of the participants' facial expressions were analysed by FaceReader 6.0. To avoid facial expression recognition bias, the 'EastAsian' model was selected to analyse Chinese faces. Meanwhile, FaceReader's estimations were operated with both non-calibration and continuous calibration models to remove the person-specific bias. Since the main findings in this study from these two models were not different,

the authors only report non-calibration model estimation results as recommended by the FaceReader manual.

Reports of FaceReader 6.0 included two files. The first one displays a temporal distribution of dominant emotions throughout the whole viewing process. An emotion is perceived dominant when its intensity is higher than that of others at the same time point. For instance, 35% of happy means in 35% of the time points examined, the intensity of happiness is higher than other emotions. The second is a log file, including not only the valence and arousal values at every 33 milliseconds, but also the intensities (between 0 and 1) of each of the six basic emotions along the time span. Self-reports also provided participants' intensity ratings of the six basic emotions on a five-point Likert scale (from 0 to 5; '0' means no rating).

To facilitate the comparison of FaceReader results and self-reported emotions on the same scale, a three-step data transformation was conducted. First, for specific emotion detection comparison, results from the two methods were transformed into truth values (0 or 1). For example, in each case, if happy was detected by FaceReader as a dominant emotion or rated by self-report, it was assigned a value of '1'; if not, a value of '0' was assigned. Second, the five levels of intensity ratings from the survey were transformed proportionately into values between '0' and '1'. That is, if a certain emotion was not rated by respondents, it was labelled as '0'. The original ratings of '1' were transformed to '0.2' and '5' to '1'. Proportional transformation rather than standardization (z-score) was implemented because the former keeps specific intensity values of emotions while the latter only demonstrates disperse levels of the emotion ratings. Maintaining specific intensity values is important because they are the basis for the third step to derive valence values for self-report. Unlike FaceReader which calculates emotional valence and arousal levels automatically, emotion valence in self-report can only be obtained by a function of 'the highest intensity of positive emotions (i.e., happy) minus the highest intensity of negative emotions (i.e., sad, angry, scared and disgusted)', same as what FaceReader did. 'Surprised' is excluded from the valence derivation because it is a mixed emotion that can be positive or negative (Ekman, 1999). Notably, the arousal level (i.e., the overall activation of participants' emotions, between 0 and 1) cannot be obtained from self-report because FaceReader calculated each individual's arousal level based on complicated algorithms of facial action units while respondents were unable to accurately evaluate their overall arousal level of six emotions.

Two comparisons were performed on the transformed data. First comparison focused on the discrete emotions identified and their intensity levels, and the second comparison

focused on valence of identified emotions. Since comparisons were informed by both qualitative and quantitative approaches (Ragin, 1981), inductive reasoning and statistical analysis were conducted. Moreover, to achieve an in-depth comprehension of the two emotional expressions, categorization of differences and theoretical interpretation were performed.

4. FINDINGS

4.1 Differences in Specific Emotions Detected by Two Methods

Differences were observed in the six basic emotions identified from FaceReader and survey. Notably, FaceReader can capture more emotional responses from the participants' faces, while self-report based on memory recalls and human perceptions revealed fewer items. As Table 2 shows, aside from the overlapped areas which are marked as grey/diamond-grid cells, the green/vertical-lined cells demonstrating emotions solely detected by FaceReader outnumber the red/horizontal-lined cells illustrating self-reported emotions only. However, even though FaceReader detected more emotions than that reported by respondents, self-report is still necessary since divergent emotions were identified from the survey and interviews. Emotions detected from these two approaches are complementary and can cross-validate each other.

Table 2.

Emotions Reported by Two Methods

Code	Happy	Sad	Angry	Surprised	Scared	Disgusted
P1_V3	green	green	green	red		grey
P1_V8			green	green		red
P1_V10		green				grey
P2_V1				green		red
P2_V7			green			grey
P3_V3		red	green		green	red
P3_V7	green				green	red
P3_V9	green	green		red	green	
P4_V4		green	green	green		
P4_V5	green		green			
P4-V8		green			green	green
P5_V5	green			red		red
P5_V6		grey	red	green	green	red
P6_V4		green	green			green
P7_V5	green		green	grey		red
P7_V8			grey	green		grey
P8_V4	green			green		grey
P8_V9	green			grey		
P9_V2	grey	green				green
P9_V3	green			red		grey
P10_V7		green		green		red
P11_V7			green			
P12_V9	grey		green	grey		
P12_V10		green	grey			red
P13_V2	red	green	green	green		green
P13_V6		grey	grey	red	red	grey
P13_V10		grey		grey	red	grey
P14_V1	green	grey	grey	grey		red
P14_V2	grey	green	green	green		

Note: Green and vertical-lined cells: identified by FaceReader only; red and horizontal-lined cells: identified by self-report survey only; grey and diamond-grid cells: identified by both.

Differences in specific emotions reported by the two methods embody in two aspects. First, different emotions were identified. As shown in Table 2, more differences (shown as 'green/vertical-lined' vs. 'red/horizontal-lined' cells) than similarities (shown as 'grey/ diamond-grid' cells) can be observed. Generally, FaceReader detected more emotions of 'happy', 'sad' and 'angry' of the respondents than self-reports did. However, self-reports expressed the emotion of 'disgusted' more frequently. For example, when participant P1 watched V10, she experienced emotional changes along the viewing process (Figure 1a), demonstrated in the fluctuation of three dominant emotions, 'sad', 'angry' and 'disgusted'. A constant low level of neutral state was also detected throughout the watching process. Figure 1b shows that during 46.8% of the viewing time, the dominant emotion was 'angry,' while 33.0% of the time was 'sad'. But in the survey, this participant only reported a feeling of 'disgusted' after watching the video. The discrepancies in specific emotions detected by the

1 two methods demonstrated distinctions between expressive channels of emotions and the
2 necessity of combining different methods for robust results.
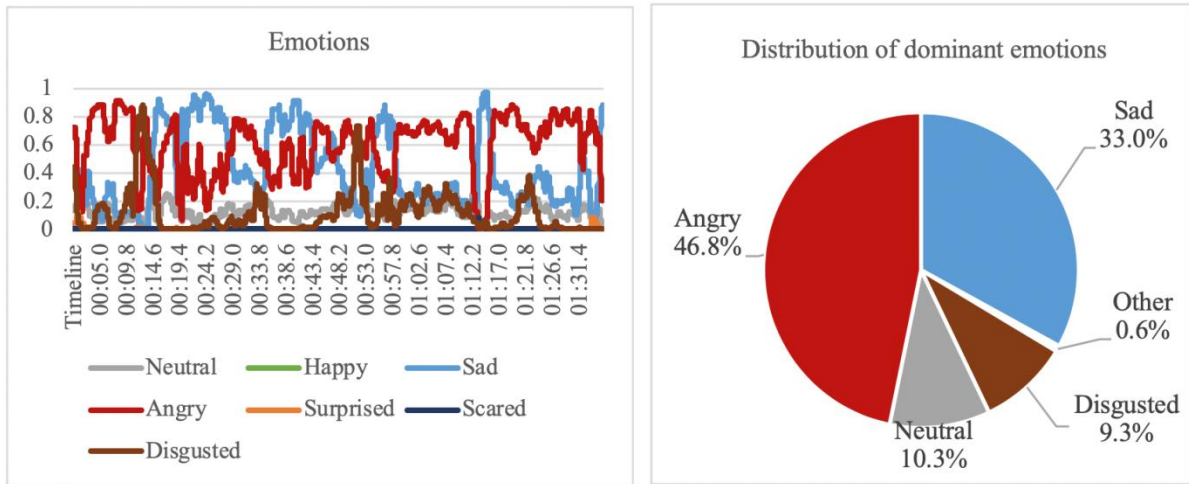
3 a. Emotional changes along timeline   b. Emotional states summary

4 *Figure 1*. **FaceReader Results of P1_V10**

5   The other aspect of differences is manifested in the intensities of specific emotions
6 reported. Although the intensities of emotions were measured in different scales by the two
7 methods, the levels of intensities can be observed within their own ranges. For instance, the
8 intensity of 'disgusted' of the case P7_V8 was identified as low (below 0.3 out of 1) by
9 FaceReader throughout the viewing process (Figure 2a). However, P7 rated the level of
10 'disgusted' highest on a 5-point Likert scale. The follow-up interview also showed a strong
11 feeling of 'disgusted': *I just felt that "it's them again"! They never make any progress. I feel*
12 *very disgusted with them'* (P7_V8). Similarly, a large difference existed in the intensities of
13 'surprised' reported by FaceReader and self-report for participant P8 when watching V9, the
14 former identified a low intensity between 0 and 0.4 (Figure 2b), whereas the latter reported a
15 high intensity of '5' in the survey. The discrepancy in intensity of emotions reflects that the
16 two methods not only differ in specific emotions identified, but also the intensity of emotions
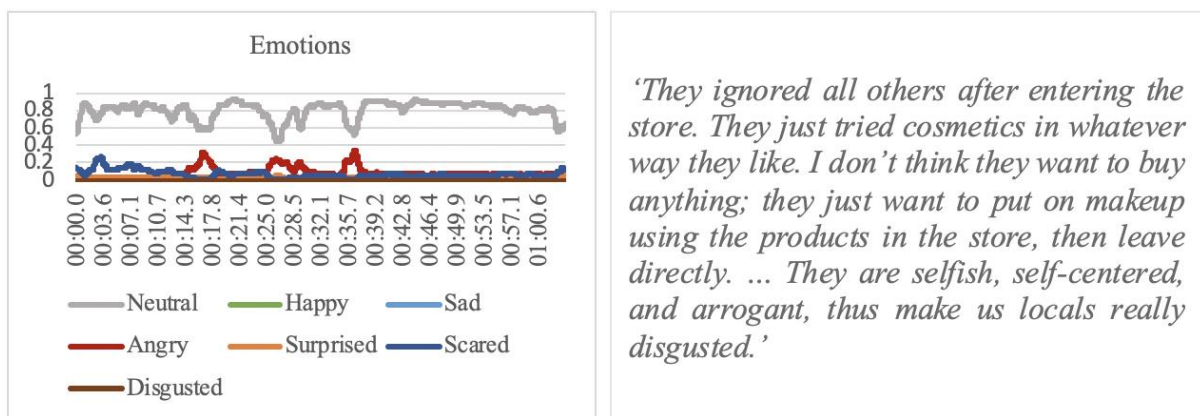17 even if the same emotion was reported.
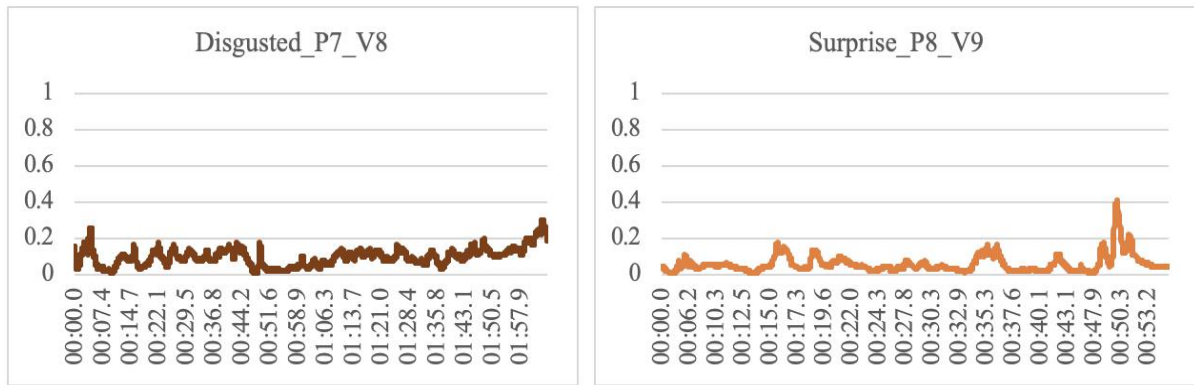
a. Intensity of 'disgusted' of P7_V8.   b. Intensity of 'surprise' of P8_V9

***Figure 2.* Examples of Emotion Intensities reported by FaceReader**

The two aspects of discrepancies in emotional responses generated by verbal reports and FaceReader can lead to contrary outcomes. For example, the facial expression recognition for the case P3_V3 revealed a 'neutral' emotional state almost over the entire viewing process, with a low intensity of all six basic emotions (lower than 0.4, see Figure 3a). However, the self-report of P3 after watching V3 expressed intensive emotions of 'sad' and 'disgusted'. The ratings of both emotions are very strong ('5'), and her discourse also reflects a strong emotion of 'disgusted' (Figure 3b). When tourists encounter this resident, they may not find her unfriendly or being offended, thereby missing the signals for improvement in their behaviours and host-tourist relations. Relying on facial expressions or self-reports alone can be misleading.

'They ignored all others after entering the store. They just tried cosmetics in whatever way they like. I don't think they want to buy anything; they just want to put on makeup using the products in the store, then leave directly. ... They are selfish, self-centered, and arrogant, thus make us locals really disgusted.'

a. Emotions of P3_V3   b. Interview of P3_V3

***Figure 3.* Emotions of P3_V3**

The differences described above are based on observations of the data. Six Mann—Whitney U tests were conducted to further examine differences in emotions detected by the

two methods. The Mann—Whitney U test was used because it is a non-parametric test that can compare the distributions of two independent samples on the condition of small sample sizes (below 30) and unnormal distributions (MacFarland & Yates, 2016).

As Table 3 illustrates, the emotions of 'happy', 'sad' and 'angry' detected by FaceReader are significantly different from that by self-report. With positive mean differences, self-report is significantly less effective in recognising happiness, sadness and anger of Hong Kong residents toward Mainland tourists. In contrast, no significant difference was identified on emotions of 'surprised', 'scared' and 'disgusted', implying a higher consistency between the two approaches when identifying these three emotions compared with the others.

Table 3.

Mann-Whitney U Test Results

FaceReader vs. Self-Report	Mean difference [a]	Mann—Whitney U value [b]
Happy	**0.310**	290.0*
Sad	**0.483**	217.5***
Angry	**0.552**	188.5***
Surprised	**0.138**	362.5
Scared	**0.103**	377.0
Disgusted	**-0.138**	478.5

Note: [a] The values for comparisons are '0' and '1'.
[b] * significant at .05 level, *** significant at .001 level.

Notably, the emotion of 'scared' was not detected much in the study (see Table 2). Compared to 'surprised' and 'disgusted' that were largely detected by both FaceReader and self-report, 'scared' was less reported by either method. This illustrates the relatively low occurrence of the 'scared' emotion in host-tourist encounters.

4.2 Differences in Emotional Valence Measured by Two Methods

Figure 4 illustrates the valence values of the two data sources across the 29 cases. For FaceReader, the average valence of emotions for each data case was derived from the real-time values. While valence values for self-reports were calculated based on intensities of specific emotions (as discussed in Methodology).

1

2 *Figure 4.* **Comparison of Emotion Valence from FaceReader and self-report Survey**

3   As shown in Figure 4, both methods identified more negative emotions because most

4 video vignettes depicted negative behaviours of Mainland tourists in Hong Kong. Overall,

5 participants' emotional valence from self-reports fluctuates more dramatically than that from

6 FaceReader results, implying that participants tend to amplify their emotions in self-reports. In

7 particular, self-reported results show higher valence values on the negative side, which means

8 participants expressed stronger negative emotions after watching videos than that could be

9 identified from their facial expressions during watching the videos. To further investigate the

10 differences in valence, a discrepancy analysis was carried out based on types of video stimuli.

11   Table 4 lists the emotion valence of the 29 cases in two groups: one with negative video

12 stimuli and the other with positive stimuli. In the group with negative stimuli, participants

13 mainly reported negative emotions with higher intensities than that detected from their facial

14 expressions. Similarly, participants reported more positive emotions for positive stimuli, also

15 with relatively higher intensities than those detected from their facial expressions. These are

16 consistent with the observation from Figure 4, suggesting the relatively amplification effect of

17 self-reports in measuring emotions in comparison with facial expressions.

18 **Table 4.**

19 *Comparison of Emotion Valence Evoked by Negative and Positive Videos*

Video clips	Case	Valence_FaceReader	Valence_Self-report	Distance (FaceReader – Self-Report)
Negative stimuli:	P2_V1	-0.78	-0.4	-0.38
	P14_V1	-0.07	-0.6	0.53
	P1_V3	0.58	-0.8	1.38

V1	P3_V3	-0.16	-1.0	0.84
V3	P9_V3	-0.27	-0.8	0.53
V4	P4_V4	-0.36	0	-0.36
V6	P6_V4	-0.35	0	-0.35
V7	P8_V4	0.31	-0.6	0.91
V8	P5_V6	-0.24	-1.0	0.76
V10	P13_V6	-0.49	-0.6	0.11
	P2_V7	-0.84	-0.4	-0.44
	P3_V7	-0.04	-1.0	0.96
	P10_V7	-0.61	-1.0	0.39
	P11_V7	-0.42	0	-0.42
	P1_V8	-0.50	-0.8	0.30
	P4_V8	-0.40	0	-0.40
	P7_V8	-0.16	-1.0	0.84
	P1_V10	-0.73	-1.0	0.27
	P12_V10	-0.49	-0.6	0.11
	P13_V10	-0.35	-0.6	0.25
	P9_V2	-0.39	0.4	-0.79
	P13_V2	-0.23	0.6	-0.83
Positive stimuli:	P14_V2	0.15	0.4	-0.25
	P4_V5	0.40	0	0.40
V2	P5_V5	-0.41	0	-0.41
V5	P7_V5	0.05	-0.6	0.65
V9	P3_V9	-0.07	0	-0.07
	P8_V9	0.46	0	0.46
	P12_V9	0.06	0.8	-0.74

The six negative stimuli cases that showed lower valence values in self-report (i.e., the distance between FaceReader and self-report survey was lower than 0) are either from participant P2 (including P2_V1 and P2_V7) or of no self-reported emotions. The inconsistent results from P2 can be due to personal facial expression bias; that is his typical facial expression tends to be negative. The other cases (i.e., P4_V4, P6_V4, P11_V7, and P4_V8) are possibly influenced by the social desirability bias in self-report. Although they did not indicate any of the six basic emotions in the survey (see red and grey cells in Table 2), these feelings were expressed during the interview after probing. For instance, P6 after watching V4 implicated a 'disgust' emotion:

> *'The video reflects an objective fact which is hideous. Because Mainland*
> *tourists always do such things; not only Hong Kong residents but also tourists*
> *from other places may have the first impression that Chinese tourists never*
> *follow the rules and like jumping a queue. This is not a good thing, very*
> *negative.'*

Likewise, positive stimuli viewers may also be subject to social desirability bias in self-report, resulting in three cases with valence distance between FaceReader and self-report

20

bigger than '0'. To further numerically verify the discrepancies in emotion valence values measured by the two methods, three non-parameter Mann—Whitney U tests were conducted. Table 5 reports a significant difference for negative stimuli. Since the negative stimuli mainly elicit negative emotions, this result implies an amplification effect of Hong Kong residents when reporting negative emotions toward Mainland tourists in the survey. The positive emotions elicited by positive stimuli are also slightly more intensive in self-reports but not statistically significant.

Table 5.

Tests of Valence Difference between FaceReader-detected and Self-reported Emotions

FaceReader vs. Self-Report	Mean difference	Mann—Whitney U value
Overall valence	0.147	333
Negative stimuli	0.292	112*
Positive stimuli	-0.176	49

Note: *significant at .05 level.

By and large, the results demonstrate that verbal self-reports amplified the negative emotions in comparison with their facial expressions. This amplification effect should be noted because it triggers antagonism in host-tourist interactions, particularly between Hong Kong residents and Mainland Chinese tourists. The emotions, especially the negative ones, read by Mainland tourists from their Hong Kong hosts' facial expressions are minor compared to the self-reported emotional responses of local residents. This disparity may make Mainland tourists underestimate the residents' unwelcome attitude or overestimate the host-tourist relation, thereby causing irreversible relationship deterioration.

5. DISCUSSION

Upon identifying the discrepancies in both specific emotions detected and the valence distances between the two methods, further analyses and theoretical interpretations are necessary to deepen the understanding of Hong Kong residents' emotional responses toward Mainland tourists.

5.1 Desires-derived vs. Stereotypes-elicited Emotions

The specific emotions detected in this study highlight the consistency of facial expressions and self-reports in identifying emotions of 'surprised', 'scared' and 'disgusted', but not in 'happy',

'sad' and 'angry'. This discrepancy indicates the existence of two types of emotions in host-tourist interaction: one can be described as desires-derived emotions and the other as stereotypes-elicited emotions. The typology can be supported by an everyday desire-belief psychology (Davidson, 1963). As Wellman and Banerjee (1991) posited, distinct emotional states were constructed and shaped by a network of mentalistic constructs, which can fall into two generic classes: desires and beliefs. Desires mean the idealised expectations toward self and others (Schroeder & Graziano, 2017). In the evolutionary history, humans have developed various desires in social interaction, such as prosocial behaviour, which become the instinctive expectations independent of human awareness (Uriely, Ram, & Malach-Pines, 2011). Accordingly, the desires-derived emotions in human interaction include 'happiness', 'sadness' and 'anger' that are elicited by satisfying (happiness) or dissatisfying (sadness/anger) humans' inner desires (Yu & Dean, 2001). Beliefs mean the cognitive impressions that are acquired through social learning and are perceptible by humans (Davidson, 1963). In social interaction, stereotypes, particularly negative ones, are widely identified beliefs in emotion elicitation (Kunda & Oleson, 1995; Schneider, 2005). Once the stereotypes are formed, stereotypes-elicited emotions, such as 'surprise', 'scare' and 'disgust' are likely to be aroused by disconfirming (surprise) or confirming (scare/disgust) people's ingrained stereotypes (Fu, 2015; Wellman & Banerjee, 1991).

The desires-derived and stereotypes-elicited emotions can be differentiated further from two aspects. First and foremost, they are initially elicited at unconscious and conscious levels, respectively, thus occurring at different ages and having distinct priorities in expressive channels (Wellman & Banerjee, 1991). Desires-derived emotions are based on humans' idealised expectations, which are typically intuitive desires stored in unconscious mind for good relationships (Uriely et al., 2011). Hence, they are available to newly-born babies and can be present in facial expressions unconsciously. For instance, infants can express emotions of happiness by smiling in mom's arms and sadness by crying when they are hungry (Ridgeway, Waters, & Kuczaj, 1985). These emotions, irrespective of positive or negative, can reflect individuals' good expectations of surrounding happenings (Nesse, 1990). On the other hand, stereotypes-elicited emotions are usually provoked by preconceived images, normally negative, in humans' conscious mind (Devine, 1989). It is not until 6-year-old that children can develop independent beliefs that may elicit 'surprised', 'scared' and 'disgusted' emotions (Fu, 2015). With a conscious elicitation, stereotypes-elicited emotions are more prevalent in self-reports. This distinction well explains the advantage of FaceReader in detecting desires-derived

emotions (i.e., 'happiness', 'sadness' and 'anger') and the effectiveness of self-report in identifying stereotypes-elicited emotions (i.e., 'surprise, 'scare' and 'disgust').

Second, the two types of emotions can be distinguished by psychological tendencies. The desires-derived emotions sourced from prosocial expectations indicate a tendency to achieve mutual beneficial relations (Schroeder & Graziano, 2017), thereby being conductive to interpersonal relationship improvement. No matter 'happy', 'sad', or 'angry', these emotions imply an expectation of Hong Kong residents to improve relationship with Mainland tourists:

> *'I hope Mainland tourists can learn from this role model (i.e., the character in*
> *this positive scenario) and improve their behaviours.'* (P3_V9)

However, when the stereotypes-elicited emotions are triggered, especially 'scared' and 'disgusted', individuals may be less inclined to actively improve relationships or situations, but would rather refuse or avoid further interactions or experience of similar events again (Lerner, Li, Valdesolo, & Kassam, 2015). Even though participant P5 reported 'surprise' after watching V5, which is a positive stimulus, his overall negative impression of Mainland tourists was not changed:

> *'The scenario is totally unlike what Mainlanders would do in reality… they*
> *definitely would not ask in that way. I have never thought he would ask, even*
> *once; but he did twice in the video! I was very surprised…But my impression of*
> *Mainland tourists was not changed… This video clip beautifies them.'*

Evidently, the stereotypes-elicited emotions do not favour the improvement of mutual relations.

Identification of the two distinct types of emotions demonstrates the ambivalent sentiment of Hong Kong residents toward Mainland Chinese tourists. Irrespective of the dominant stereotypes-elicited emotions in respondents' self-reports, Hong Kong residents also have hidden positive expectations of Mainland tourists, as shown in the desires-derived emotions recognised from their facial expressions when watching the videos. It is the long-term dissatisfaction with Mainland tourists whose behaviours never meet their expectations caused deep-rooted Hong Kong residents' disgusted or resentful feelings. Hong Kong locals become increasingly impatient or hopeless to see any improvements even though deep down they wish this could be the case (in Chinese, 恨铁不成钢). Driven by this ambivalent sentiment, the respondents disclosed their stereotypes-elicited emotions in self-reports but hid or failed to be aware of their concealed expectations.
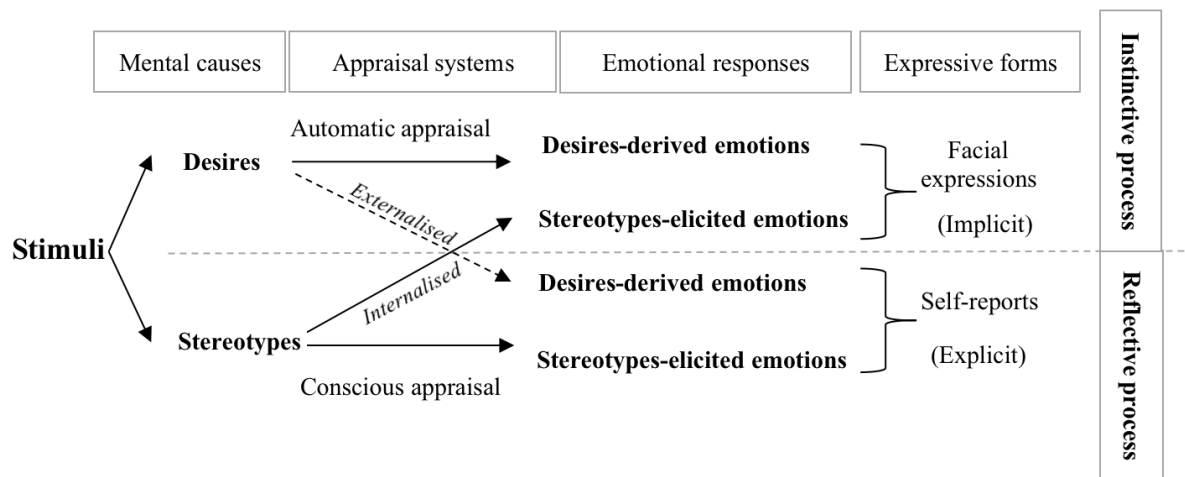
5.2 Facial Expressions vs. Words: Unlocking Emotion Formation Processes

The observed and numerically verified discrepancies in emotional responses of Hong Kong residents toward Mainland tourists measured by FaceReader and self-report show that implicit and explicit expressions of emotion deliver different messages. Self-reported emotions as explicit expressions have bias both in specific items and their valence. Participants reported more emotions of 'surprise', 'scare' and 'disgust' in the interview and survey, which are mainly aroused by their preconceived stereotypes, but less frequently reported the more instinctive, desires-derived emotions, such as 'happiness,' 'sadness' and 'anger'. Moreover, the respondents tend to amplify the intensity of negative emotions (i.e., 'sad', 'angry', 'scared' and 'disgusted') in self-reports. With the low occurrence of 'scared' in host-tourist interactions (Faullant et al., 2011), 'disgusted' thus became the dominant emotion in self-reports within this research setting of Hong Kong residents–Mainland tourists interaction:

> *'Firstly I felt funny because it is very realistic. Then I felt disgusted; their*
> *behaviours have exerted bad influences on our daily lives. Such behaviours are*
> *really annoying.'* (P12_V10)

However, FaceReader identified that participants also held desires-derived emotions in an implicit form, demonstrating good expectations of Hong Kong residents toward Mainland tourists. These complex emotional experiences and connections between hosts and tourists could not be fully captured without considering implicit and explicit emotional expressions together.

The empirically identified discrepancies in this study, including different mental causes (i.e., desires vs. stereotypes) and their corresponding emotions (i.e., desires-derived vs. stereotypes-elicited) expressed in different forms (i.e., implicit vs. explicit), can be connected to the two appraisal systems to support a dual-process of emotion formation: an instinctive process underlying the automatic appraisal and a reflective process underlying the conscious appraisal (see Figure 5). The two processes of emotion formation were named in accordance with the dual-system of appraisal because appraisal plays a distinguished role in emotion formation and expression (Haidt, 2001). The dual-process model of emotion formation proposed by this study contributes to the literature by specifying the discrepancies and the underlying psychological mechanisms that determine distinct emotional outcomes in different expressive forms. Desires and stereotypes are the main drivers respectively in the two processes of emotion formation.

The diagram contains the following labeled boxes and text:

Mental causes | Appraisal systems | Emotional responses | Expressive forms

Instinctive process

Stimuli

Desires — Automatic appraisal → Desires-derived emotions

Externalised

Stereotypes-elicited emotions → Facial expressions (Implicit)

Internalised

Desires-derived emotions

Stereotypes — Conscious appraisal → Stereotypes-elicited emotions → Self-reports (Explicit)

Reflective process

1

Figure 5. **Dual-process Model of Emotion Formation**

3   The instinctive process creates immediate emotional responses to the stimuli based on
4 the activation of unconscious or innate mental associations (Westen, 1999). Emotions
5 generated from this process are usually expressed implicitly (Haidt, 2001), such as in facial
6 expressions with an automatic and rapid onset, and with a short duration. Among the many
7 other instinct and primitive associations, the idealised expectation (i.e., desires) of humans,
8 such as the pro-sociality, is an important realm that can elicit desires-consistent emotions
9 outside of human awareness (Kahneman & Frederick, 2001). Moreover, long-term existence
10 or continuously strengthening of stereotypes may lead to their internalisation, so that conscious
11 associations become unconscious (Tran et al., 2011) and stereotypes-elicited emotions can also
12 be aroused subliminally. Taking two cases as examples, when P1 watched V10, apart from
13 'sad' and 'angry' emotions elicited against her expectations, the 'disgusted' emotion was also
14 identified in her facial expressions because of her preconceived stereotypes:

15   *'It is very common for Mainland tourists to pack suitcases on the street. I've*
16   *seen similar scenes too many times before.'*

17 Similarly, with a positive V5, 'surprised' was identified in P7's facial expressions because what
18 the video showed was totally different from her ingrained negative stereotypes:

19   *'The characters in the video are not as unruly as those Mainland tourists I*
20   *met before.'*

21   The reflective process involves a more conscious processing of perceived stimuli,
22 which activates the cognitive mental associations and leads to explicit expression of emotions
23 (Westen, 1999) that can be measured by retrospective self-reports. The cognitive associations,

at the first place, relate to the stereotypes in humans' mind which refer to the knowledge and beliefs learned by reflective thinking (Tung, King, & Tse, 2019). They can quickly affect individuals' cognitive appraisals and lead to stereotypes-consistent emotions (Cuddy, Fiske, & Glick, 2007). In this study, respondents felt 'surprised' when the scenarios disconfirmed their stereotypes of Mainlanders but felt 'disgusted' when the negative stereotypes are confirmed. For instance, participant P7 reported 'surprised' when watching Mainland tourists behave politely in V5 by stating that it would be normal if the characters were Japanese or Caucasians. The distinct emotional responses to tourists of different origins confirm the critical role of stereotypes in eliciting emotions through the reflective process.

Stereotypes play a dominant role in the reflective process because of their origin in reflective learning and conscious thinking (Westen, 1999). However, as shown in Table 2, desires-derived emotions of 'happy', 'sad' and 'angry', though not common, have also been reported explicitly. The reason may be due to instinctive desires can be externalised (Uriely et al., 2011) to exert influence on cognitive judgements and arouse corresponding emotions. For instance, while participant P5 reported 'disgusted' after watching V6, he perceived 'sad' and 'angry' as well because the scenario in the videoclip did not meet his positive expectations of Mainland tourists:

'I did not expect them to behave like this. This couple were rude and out of my expectation. I thought they could be better. But the reality is not.'

The influence of desires on conscious appraisal and the subsequent emotional responses are presented with dotted lines in Figure 5 due to a lesser extent of impacts compared to stereotypes at the conscious level (i.e., significantly less reports of desires-derived emotions in survey, but no significant difference in stereotypes-elicited emotions was observed between FaceReader and self-report results).

Hong Kong residents experienced desires-derived emotions and stereotypes-elicited emotions toward Mainland Chinese tourists but expressed them differently – some reserved in facial expressions and some amplified in words – because of the two distinct emotion formation processes. This study illustrated that the long-term negative stereotypes of Mainland Chinese tourists has been internalised deep into Hong Kong residents' mind to exert influences on emotional responses within both processes. Furthermore, the internalised stereotypes of Mainland tourists in the mind of Hong Kong residents can, to some extent, deter the externalization of desires and explicit expression of desires-derived emotions. As a result,

Hong Kong residents tended to report the emotions consistent with their negative stereotypes and in particular, amplify the 'disgusted' feeling in self-reports.

6. CONCLUSION

Emotion is a long-standing but still deserving topic owing to its uncertainty and complexity in expressions. In social interactions, particularly the host-tourist interaction in international tourism, emotion plays a critical role in determining host-tourist relations (Lerner et al., 2015). People express emotions through various implicit and explicit channels, including face and body gestures, words, and voice (Baumeister & Bushman, 2020). However, emotion research in tourism, even in marketing and psychology, has a long tradition of relying on self-report method to examine emotional responses expressed in explicit forms, but ignoring their implicit expressions (Prayag, 2020; Scherer, 2005a). Using FaceReader, the study compared emotional responses of Hong Kong residents toward Mainland Chinese tourists detected from their real-time facial expressions with their self-reported emotions to understand the complexity of emotions in a host-tourist interaction context.

The discrepancies identified in emotional responses measured by the two methods demonstrate that facial expressions and words can tell different stories. Specifically, facial expression recognition is better at identifying desires-derived emotions including happiness, sadness, and anger, while self-reports tend to amplify the stereotypes-elicited and negative emotions, in particular 'disgust'. These two types of emotions (desires-derived vs. stereotypes-elicited) we identified demonstrated the ambivalent sentiments (恨铁不成钢) that Hong Kong residents hold toward Mainland tourists. To further explain the discrepancies in facial expressions and words, a dual-process model of emotion formation was proposed to illustrate the specific psychological mechanisms that determine the different emotional outcomes in distinct expression forms. The instinctive process illuminates the formation of emotions that aroused by both desires and implicit stereotypes, while the reflective process accounts for the formation of explicit expressions of emotions elicited by stereotypes and externalised desires.

This empirical study contributes to the theoretical interpretation of emotional discrepancies in different expressive forms and offers innovative insights, through facial expressions of residents, on emotion research in tourism (Cohen & Cohen, 2019). Different from previous studies that revealed mainly negative stereotypes of Hong Kong residents toward Mainland Chinese tourists (Chen et al., 2018; Tung et al., 2019), this study identified the

coexistence of stereotypes and good expectations in the mind of Hong Kong residents. Different emotions imply different behavioural tendencies (Fu, 2015). This ambivalent sentiment of residents toward tourists deserves further exploration for practical improvements of host-tourist relations. As to this study context, the Hong Kong Tourism Board/Commission and local media should promote a better image of Mainland tourists to gradually modify local residents' negative stereotypes and reduce stereotypes-elicited emotions. For tourists, aside from using facial expressions as a main channel to understand hosts' emotional responses, more attention needs to be paid on residents' verbal expressions. Comprehending each other's emotional responses from multiple channels is necessary for a friendly inter-personal/group relationship. Mainland tourists could also improve their behaviours in Hong Kong by respecting local rules and customs, to meet local residents' expectations and foster desires-derived emotions to benefit mutual relations.

This study focused on the context of Hong Kong resident-Mainland Chinese tourist interaction; specific results may not be generalized to other host-tourist interactions due to the special historical relation between these two groups. Future studies should be conducted in other destinations to attest the existence of two different types of emotions in residents' responses toward tourists, and further refine the dual-process model of emotion formation. Although the personal interviews reached data saturation, this study identified emotional expressive discrepancies on a relatively small sample size in terms of the statistically significance. Further studies on a larger sample to verify the two types of emotional responses and test their connections to the two mental causes (desires and stereotypes) and expressive forms are encouraged. Other influential factors of resident emotional responses toward tourists, such as demographics, identification with tourists, personal experience with tourism, and duration differences among various emotions (Verduyn & Lavrijsen, 2015) should also be considered in future studies. Additionally, the overwhelming influence of stereotypes on explicit emotional responses as well as host-tourist interaction quality deserves further investigation. The video vignettes used in this study could be re-utilized to explore tourist meta-stereotypes and its influence on tourist behaviours, which is an emerging research stream in both social psychology (e.g., Rodriguez Mosquera, Uskul, & Cross, 2011) and tourism (e.g., Tung, 2019). Finally, other relevant theories such as social identity theory and theories of emotion causation can be applied to advance interpretations of complex emotional responses. To further these research, combination of various psychophysiological measures (e.g., EEG, EDA and Facial expressions) and self-reports are strongly encouraged.

REFERENCES

Arnold, M. B. (1960). *Emotion and personality*. London: Cassell.

Baumeister, R. F., & Bushman, B. J. (2020). *Social Psychology and Human Nature* (5 ed.). Boston: Cengage Learning.

Boz, H., & Koc, E. (2019). Service quality, emotion recognition, emotional intelligence and Dunning Kruger syndrome. *Total Quality Management & Business Excellence*, 1-14.

Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on affective computing, 1*(1), 18-37.

Caramani, D. (2008). *Introduction to the comparative method with Boolean algebra* (Vol. 158). Newbury Park: Sage publications.

Chen, N., Hsu, C. H., & Li, X. R. (2018). Feeling superior or deprived? Attitudes and underlying mentalities of residents towards Mainland Chinese tourists. *Tourism Management, 66*, 94-107.

Cheung, K. S., & Li, L.-H. (2019). Understanding visitor–resident relations in overtourism: developing resilience for sustainable tourism. *Journal of Sustainable Tourism, In press*, 1-20.

Cohen, S. A., & Cohen, E. (2019). New directions in the sociology of tourism. *Current Issues in Tourism, 22*(2), 153-172.

Crivelli, C., & Fridlund, A. J. (2019). Inside-out: From basic emotions theory to the behavioral ecology view. *Journal of Nonverbal Behavior, 43*(2), 161-194.

Cuddy, A. J., Fiske, S. T., & Glick, P. (2007). The BIAS map: behaviors from intergroup affect and stereotypes. *Journal of personality and social psychology, 92*(4), 631.

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

Davidson, D. (1963). Actions, reasons, and causes. *The Journal of Philosophy, 60*(23), 685-700.

De Houwer, J. (2006). What are implicit measures and why are we using them. In R. W. Wiers & A. W. Stacy (Eds.), *The handbook of implicit cognition and addiction* (pp. 11-28). London: Sage Publications.

Denzin, N. K., & Lincoln, Y. S. (2008). *The landscape of qualitative research* (Vol. 1): Sage.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of personality and social psychology, 56*(1), 5.

Drozdova, N. (2014). *Measuring Emotions in Marketing and Consumer Behavior: Is Face Reader an applicable tool?* (Master), Norwegian School of Economics, Bergen.

Ekman, P. (1970). Universal facial expressions of emotion. *California Mental Health Reserch Digest, 8*(4), 151-158.

Ekman, P. (1985/2009). *Telling lies: Clues to deceit in the marketplace, politics, and marriage (revised edition)*. New York, London: WW Norton & Company.

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. (1999). Basic emotions. *Handbook of cognition and emotion, 98*(45-60), 16.

Esposito, A. (2009). Affect in multimodal information. In J. Tao & T. Tan (Eds.), *Affective Information Processing* (pp. 203-226). London: Springer.

Faullant, R., Matzler, K., & Mooradian, T. A. (2011). Personality, basic emotions, and satisfaction: Primary emotions in the mountaineering experience. *Tourism Management, 32*(6), 1423-1430.

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

Fu, Y. (2015). *Emotional psychology*. Shanghai: East China Normal University Press.

Gakhal, B., & Senior, C. (2008). Examining the influence of fame in the presence of beauty: An electrodermal 'neuromarketing'study. *Journal of Consumer Behaviour: An International Research Review, 7*(4-5), 331-341.

González-Rodríguez, M. R., Díaz-Fernández, M. C., & Gómez, C. P. (2020). Facial-expression recognition: An emergent approach to the measurement of tourist satisfaction through emotions. *Telematics and Informatics*, 101404.

Hadinejad, A., Moyle, B. D., Kralj, A., & Scott, N. (2019). Physiological and self-report methods to the measurement of emotion in tourism. *Tourism Recreation Research, In press*, 1-13.

Hadinejad, A., Moyle, B. D., Scott, N., & Kralj, A. (2019). Emotional responses to tourism advertisements: the application of FaceReader™. *Tourism Recreation Research, 44*(1), 131-135.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological review, 108*(4), 814.

Hetland, A., Vittersø, J., Fagermo, K., Øvervoll, M., & Dahl, T. I. (2016). Visual excitement: analyzing the effects of three Norwegian tourism films on emotions and behavioral intentions. *Scandinavian Journal of Hospitality and Tourism, 16*(4), 528-547.

Hosany, S., Martin, D., & Woodside, A. G. (2020). Emotions in Tourism: Theoretical Designs, Measurements, Analytics, and Interpretations. *Journal of Travel Research* (In press), 1-17. doi:https://doi.org/10.1177/004728752093707

Izard, C. E. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.

Jordan, E. J., Spencer, D. M., & Prayag, G. (2019). Tourism impacts, emotions and stress. *Annals of Tourism Research, 75*, 213-226.

Kahneman, D., & Frederick, S. (2001). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and biases: The psychology of intuitive judgment, 49*, 81.

Keltner, D., Tracy, J. L., Sauter, D., & Cowen, A. (2019). What basic emotion theory really says for the twenty-first century study of emotion. *Journal of Nonverbal Behavior, 43*(2), 195-201.

Kim, J., & Fesenmaier, D. R. (2015). Measuring emotions in real time: Implications for tourism experience design. *Journal of Travel Research, 54*(4), 419-429.

Koc, E., & Boz, H. (2020). Development of hospitality and tourism employees' emotional intelligence through developing their emotion recognition abilities. *Journal of hospitality marketing & management, 29*(2), 121-138.

Konstan, D. (2006). *The emotions of the ancient Greeks: Studies in Aristotle and classical literature* (Vol. 5). Toronto: University of Toronto Press.

Kunda, Z., & Oleson, K. C. (1995). Maintaining stereotypes in the face of disconfirmation: Constructing grounds for subtyping deviants. *Journal of personality and social psychology, 68*(4), 565.

Lagast, S., Gellynck, X., Schouteten, J., De Herdt, V., & De Steur, H. (2017). Consumers' emotions elicited by food: A systematic review of explicit and implicit methods. *Trends in food science & technology, 69*, 172-189.

Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford: Oxford University Press.

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual review of psychology, 66*, 799-823.

Lewinski, P., den Uyl, T. M., & Butler, C. (2014). Automated facial coding: Validation of basic emotions and FACS AUs in FaceReader. *Journal of Neuroscience, Psychology, and Economics, 7*(4), 227.

Li, S. (2019). Emotional Appeals In Tourism Tv Commercials: A Psycho-Physiological Study. *Journal of Hospitality & Tourism Research*, 1096348019828440.

Li, S., Scott, N., & Walters, G. (2015). Current and potential methods for measuring emotion in tourism experiences: A review. *Current Issues in Tourism, 18*(9), 805-827.

Li, S., Walters, G., Packer, J., & Scott, N. (2018a). A comparative analysis of self-report and psychophysiological measures of emotion in the context of tourism advertising. *Journal of Travel Research, 57*(8), 1078-1092.

Li, S., Walters, G., Packer, J., & Scott, N. (2018b). Using skin conductance and facial electromyography to measure emotional responses to tourism advertising. *Current Issues in Tourism, 21*(15), 1761-1783.

Lin, C.-Y., & Lin, J.-S. C. (2017). The influence of service employees' nonverbal communication on customer-employee rapport in the service encounter. *Journal of Service Management*.

Liu, H. (2018). Comparative research method in social science: development, typology and arguments. *Social Science Abroad, 5*(1), 122-133 (in Chinese).

MacFarland, T. W., & Yates, J. M. (2016). Mann–whitney u test *Introduction to nonparametric statistics for the biological sciences using R* (pp. 103-132): Springer.

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and emotion, 23*(2), 209-237.

Moors, A. (2009). Theories of emotion causation: A review. *Cognition and emotion, 23*(4), 625-662.

Moyle, B. D., Moyle, C.-l., Bec, A., & Scott, N. (2019). The next frontier in tourism emotion research. *Current Issues in Tourism, 22*(12), 1393-1399.

Nawijn, J., & Biran, A. (2019). Negative emotions in tourism: A meaningful analysis. *Current Issues in Tourism, 22*(19), 2386-2398.

Nesse, R. M. (1990). Evolutionary explanations of emotions. *Human nature, 1*(3), 261-289.

Nilsen, L., & Kaszniak, A. (2007). Conceptual, theoretical and methodological issues in inferring subjective emotion Experience. In J. A. Coan & J. J. Allen (Eds.), *Handbook of emotion elicitation and assessment* (pp. 361-379). Oxford: Oxford University Press.

Penn, D. (2006). Looking for the emotional unconscious in advertising. *International Journal of Market Research, 48*(5), 515-524.

Pestana, M. H., Parreira, A., & Moutinho, L. (2019). Motivations, emotions and satisfaction: The keys to a tourism destination choice. *Journal of Destination Marketing & Management, In press*.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In R. Plutchik & H. Kellerman (Eds.), *Theories of emotion* (pp. 3-33). New York: Academic Press.

Prayag, G. (2020). The role and measurement of emotion in tourism experiences. In S. K. Dixit (Ed.), *The Routledge Handbook of Tourism Experience Management and Marketing* (pp. 77-87). New York: Routledge.

Prayag, G., Hosany, S., Muskat, B., & Del Chiappa, G. (2017). Understanding the relationships between tourists' emotional experiences, perceived overall image, satisfaction, and intention to recommend. *Journal of Travel Research, 56*(1), 41-54.

Ragin, C. C. (1981). Comparative sociology and the comparative method. *International Journal of comparative sociology, 22*(1-2), 102-120.

Ridgeway, D., Waters, E., & Kuczaj, S. A. (1985). Acquisition of emotion-descriptive language: Receptive and productive vocabulary norms for ages 18 months to 6 years. *Developmental Psychology, 21*(5), 901.

Robinson, J. (2005). *Deeper than reason: Emotion and its role in literature, music, and art*. London: Oxford University Press on Demand.

Rodriguez Mosquera, P. M., Uskul, A. K., & Cross, S. E. (2011). The centrality of social image in social psychology. *European journal of social psychology, 41*(4), 403-410.

Sacharin, V., Schlegel, K., & Scherer, K. R. (2012). *Geneva emotion wheel rating study*. Retrieved from Geneva, Switzerland:

Sato, W., Hyniewska, S., Minemoto, K., & Yoshikawa, S. (2019). Facial Expressions of Basic Emotions in Japanese Laypeople. *Frontiers in psychology, 10*, 259.

Scherer, K. R. (2005a). Unconscious Processes in Emotion: The Bulk of the Iceberg. In L. F. Barrett, P. M. Niedenthal, & P. Winkielman (Eds.), *Emotion and consciousness* (pp. 312-334). New York: Guilford Press.

Scherer, K. R. (2005b). What are emotions? And how can they be measured? *Social science information, 44*(4), 695-729.

Scherer, K. R., & Moors, A. (2019). The emotion process: event appraisal and component differentiation. *Annual review of psychology, 70*, 719-745.

Schneider, D. J. (2005). *The psychology of stereotyping*: Guilford Press.

Schroeder, D. A., & Graziano, W. G. (2017). Prosocial behavior as a human essence *The Oxford Handbook of the Human Essence*: Oxford University Press.

Shen, H., Luo, J., & Zhao, A. (2017). The sustainable tourism development in Hong Kong: An analysis of Hong Kong residents' attitude towards mainland Chinese tourist. *Journal of Quality Assurance in Hospitality & Tourism, 18*(1), 45-68.

Smelser, N. J. (2013). *Comparative methods in the social sciences*. New Orleans: Quid Pro Books.

Somervuori, O., & Ravaja, N. (2013). Purchase behavior and psychophysiological responses to different price levels. *Psychology & Marketing, 30*(6), 479-489.

Swanson, G. (1971). Frameworks for comparative research: Structural anthropology and the theory of action. In D. E. Apter (Ed.), *Comparative methods in sociology: Essays on trends and applications* (pp. 141-202). California: Univ of California Press.

Tomkins, S. (1963). *Affect imagery consciousness: Volume II: The negative affects*. New York: Springer Publishing Company.

Tran, T. B., Siemer, M., & Joormann, J. (2011). Implicit interpretation biases affect emotional vulnerability: A training study. *Cognition and emotion, 25*(3), 546-558.

Tung, V. W. S. (2019). Helping a Lost Tourist: The Effects of Metastereotypes on Resident Prosocial Behaviors. *Journal of Travel Research, 58*(5), 837-848.

Tung, V. W. S., King, B. E. M., & Tse, S. (2019). The Tourist Stereotype Model: Positive and Negative Dimensions. *Journal of Travel Research, In press*, 0047287518821739.

Uriely, N., Ram, Y., & Malach-Pines, A. (2011). Psychoanalytic sociology of deviant tourist behavior. *Annals of Tourism Research, 38*(3), 1051-1069.

van Zonneveld, L., de Sonneville, L., Van Goozen, S., & Swaab, H. (2019). Recognition of facial emotion and affective prosody in children at high risk of criminal behavior. *Journal of the International Neuropsychological Society, 25*(1), 57-64.

Verduyn, P., & Lavrijsen, S. (2015). Which emotions last longest and why: The role of event importance and rumination. *Motivation and Emotion, 39*(1), 119-127.

Walsh, A. M., Duncan, S. E., Bell, M. A., O'Keefe, S. F., & Gallagher, D. L. (2017). Breakfast meals and emotions: Implicit and explicit assessment of the visual experience. *Journal of Sensory Studies, 32*(3), 1-13.

Wellman, H. M., & Banerjee, M. (1991). Mind and emotion: Children's understanding of the emotional consequences of beliefs and desires. *British Journal of Developmental Psychology, 9*(2), 191-214.

Westen, D. (1999). The scientific status of unconscious processes: Is Freud really dead? *Journal of the American Psychoanalytic Association, 47*(4), 1061-1106.

Wolf, K. (2015). Measuring facial expression of emotion. *Dialogues in clinical neuroscience, 17*(4), 457.

1   Yu, Y.-T., & Dean, A. (2001). The contribution of emotional satisfaction to consumer
2   loyalty. *International journal of service industry management, 12*(3), 234-250.
3   Zheng, D., Ritchie, B. W., Benckendorff, P. J., & Bao, J. (2019). Emotional responses toward
4   Tourism Performing Arts Development: A comparison of urban and rural residents in
5   China. *Tourism Management, 70*, 238-249.
6
7