Emotions in tourism:

Theoretical designs, measurements, analytics, and interpretations

Sameer Hosany, School of Business and Management, Royal Holloway University of London, UK, sameer.hosany@rhul.ac.uk

Drew Martin, School of Hotel, Restaurant and Tourism Management, University of South Carolina, USA, martin11@mailbox.sc.edu

Arch G. Woodside, School of Marketing, Curtin University, Perth, Australia, arch.woodside@curtin.edu.au and arch.woodside@bc.edu

May 2020

[Forthcoming. This a post-print version of the manuscript accepted for publication in Journal of Travel Research. For full citation, refer to the published available using the DOI:10.1177/0047287520937079]
Abstract

The theorization of emotion receives considerable attention in contemporary tourism literature. Remarkably, existing studies largely ignore the operationalization of emotion in tourism research. Drawing on extant knowledge from psychology, marketing, and tourism literatures, this article describes methodological and theoretical concerns and provides guidance for selecting highly useful-for-the-context (HUFTC) emotion measures. To help researchers choose HUFTC measures, this study proposes a new model: Emotionapps. The article here highlights the need for tourism researchers to account for the complexities in measuring emotions and how such measurement impacts theory construction.

Keywords: emotions; cognitive appraisal theories, emotionapps model; measurement; self-report
Introduction

Emotions are core to tourism experiences (Aho 2001; Bastiaansen et al. 2019; Knobloch, Robertson, and Aitken 2017; Tussyadiah 2014). Prior studies establish the relevance of emotions across various settings such as festivals (e.g., Lee 2014), shopping (Yuksel and Yuksel 2007), theme parks (Bigné, Andreu, and Gnoth 2005), holidays (Hosany and Prayag 2013), heritage sites (Prayag, Hosany, and Odeh 2013), scenic tourist attractions (e.g., Wang and Lyu 2019), and adventure tourism (Faullant, Matzler, and Mooradian 2011). Emotions influence various stages of the tourist experience. At the pre-travel stage, emotions activate tourists’ motivations and inputs in destination choice processes (Gnoth, 1997). During the trip, emotions vary on a day-to-day basis (Nawijn et al. 2013). Tourists’ emotional reactions are fundamental precursors of satisfaction (e.g., Faullant et al. 2011; Hosany et al. 2017), destination attachment (Yuksel, Yuksel, and Bilim 2010), behavioural intentions (Yuksel and Yuksel, 2007; Prayag, Hosany, and Odeh 2013) and perceived overall image evaluations (Prayag et al. 2017). In addition, an emerging body of research focuses on residents’ emotional responses toward tourism development, tourism impacts and support (e.g., Jordan, Spencer, and Prayag 2019; Ouyang, Gursoy, and Sharma 2017; Zheng et al. 2019a). Residents’ experienced emotions are determinants of support for hosting a mega event (Ouyang, Gursoy, and Sharma 2017) and tourism related stress (Jordan, Spencer, and Prayag 2019).

Travel and tourism research draw heavily on emotion measures from the psychology literature. For example, studies apply self-report emotion scales (e.g., Izard 1977; Mehrabian and Russell 1974; Plutchik 1980; Watson, Clark, and Tellegen 1988) to understand tourist experiences. Yet, despite these strong foundations, researchers question the applicability, reliability, and validity of psychological emotion measures in tourism studies (Hosany and Gilbert 2010; Lee and Kyle 2013). For example, Izard’s (1977) “Differential Emotions Scale”
(DES) focusing on facial expressions of emotion, overrepresent negative emotions. Hosany and Gilbert (2010) conclude that existing psychology-based emotion measures are context specific and fail to capture tourists’ and destinations’ specific characteristics. Method applied informs thinking, theory crafting, and theory testing (Gigerenzer 1991). Correct emotion measures have important implications for understanding tourism experiences. Until recently, the tourism literature overlooks measurement issues relating to the operationalisation of emotion. To address this gap, integrating extant knowledge from the psychology, marketing and tourism literatures, this article discusses current methods and concerns as well as provides guidance to select “highly useful-for-the-context” (HUFTC) measures.

In particular, the present study addresses the following questions. Are summary dimensions of emotion appropriate? Are adapted self-report emotion measures from psychology appropriate? When should verbal, non-verbal, or indirect qualitative emotion measures be employed? Should tourism researchers use unipolar or bipolar scales to measure emotion? Should tourism researchers capture retrospective or in-situ process emotions? How should the interplay of emotion and cognition be used in tourist behaviour models? How to apply configurational theoretical and measurement design in emotion research? To help researchers choose HUFTC measures, this paper proposes a new model: “Emotionapps”. The Emotionapps model triangulates degrees of salience, valence, and consciousness.

Conceptual papers are relatively rare in tourism research (Baum et al. 2016; Xin, Tribe, and Chambers 2013). Gilson and Goldberg (2016) note the need for good conceptual articles to integrate works, theories across disciplines and broaden the scope of our thinking. Our paper contributes to theoretical advancement by summarising, integrating and structuring extant knowledge across multiple literature streams. Theoretical and measurement issues relating to emotions appear in places but in piecemeal fashion in tourism (e.g, Kim and Fesenmaier 2015; Li et al. 2018a). The present study brings together the fragmented
literature, identify common grounds and propose an enhanced conceptualisation of emotion. Our article offers a solid starting point for readers by providing a state-of-the-art treatise of the various theoretical and methodological considerations affecting emotion research in tourism.

**Methodological-theoretical design considerations**

This study draws on relevant empirical and conceptual studies published in psychology, marketing and tourism disciplines. A domain-based (Palmatier, Houston, and Hulland 2018) synthesis approach (Jaakkola 2020) was chosen to summarise and integrate current understanding across multiple theoretical and methodological perspectives (MacInnis 2011). Such synthesis is relevant particularly when the topic of interest is fragmented across different literatures, helping to identify and highlight commonalities that build coherence (Cropanzano 2009) and points the way forward for scholars. We searched for papers in electronic databases (including Scopus, Web of Science and EBSCO) that broadly address theoretical and measurement issues relating to emotions in psychology, marketing and tourism. The method as described by Greenhalgh et al. (2005) that involves search, mapping, appraisal and evaluation, was adopted to identify relevant papers. The initial corpus of papers was complemented by conducting backward (citations in key papers identified during the initial phase were reviewed for relevance) and forward searches (papers citing key papers were reviewed) using Google Scholar (Bandara et al. 2015; Caruelle et al. 2019; Snyder, 2019). As a result of this process, 194 articles were identified across the three disciplines: psychology (N=70), marketing (N=76), and tourism (N=48). The review uncovers seven methodological and theoretical considerations in emotion research: appropriateness of summary dimensions of emotion; appropriateness of adapted self-report emotion measures from psychology; verbal, non-verbal and indirect qualitative emotion measures; unipolar versus bipolar emotions scales; the interplay of emotion and cognition in consumer behaviour
models; and linear versus configurational theory construction. The rest of the paper addresses each consideration and provides recommendations for researchers.

*Appropriateness of summary dimensions of emotion*

The psychology literature offers three main theoretical branches describing emotions: dimensional, categorical, and cognitive appraisals. Dimensional (valence-based) conceptualises emotions using a few dimensions such as positive and negative (Watson, Clark, and Tellegen 1988) or pleasantness and arousal (Russell 1980). This approach does not require distinguishing between distinct negative and positive emotions (Rucker and Petty 2004). The dimensional approach offers a parsimonious account of emotional experiences (Lazarus 1991). Empirical studies employing factor analysis often show two-dimensional structures (positive and negative) and capture a large portion of an emotional rating’s variance (Bagozzi, Gopinath, and Nyer 1999; Watson and Tellegen 1985). Unfortunately, a dimensional approach fails to specify whether or not distinct emotions of the same valence (e.g., sadness, anger, and fear, or elation and contentment) show different action tendencies (Frijda 1986; Lerner and Keltner 2000). A valence-based approach sacrifices specificity for parsimony (Higgins 1997). Mehrabian and Russell (1974) and Watson et al. (1988) are examples of common dimensional theories.

Emotions represent distinct mental states (e.g., joy, anger, or fear). Examining a few global dimensions (e.g., positive and negative) oversimplifies an emotional experience’s complexity (Bagozzi et al. 2000; Rucker and Petty 2004). Machleit and Eroglu (2000) note that combining emotional responses into summary dimensions hide relationships between specific emotions and satisfaction. A categorical approach (e.g., Izard 1977; Plutchik 1980) conceptualises emotions as a set of idiosyncratic affective states and offers a solution to exploring behavioural consequences of specific emotions. For example, emotions of the same valence (e.g., fear and anger; sadness and anxiety; regret and disappointment) differently
affect judgement (see Lerner and Keltner 2000), decision-making (Raghunathan and Phan 1999), product preference (Rucker and Petty 2004), satisfaction and complaining behaviour (Zeelenberg and Pieters 2004).

Evolution in psychology research unifies emotions’ studies in cognitive appraisal theories. Arnold (1960) coins the expression, “emotional appraisal” referring to the cognitive process involving emotion elicitation. Appraisal theories address earlier approaches’ limitations to conceptualise emotions. For example, the categorical approach does not determine the emotional causes and the dimensional approach cannot distinguish between emotions of similar valence. Appraisal theories consider cognition an antecedent of emotion. Emotions are mental states resulting from processing or appraising personally relevant information. “[E]valuations and interpretations of events, rather than events per se determine whether an emotion will be felt and which emotion it will be” (Roseman, Spindel, and Jose 1990, p. 899). For example, an event’s appraisal as beneficial and within reach elicits joy.

Different appraisal situations elicit diverse emotional reactions (e.g., Ruth, Brunel, and Otnes 2002) and lead to dissimilar behavioural consequences (e.g., Frijda and Zeelenberg 2001).

Consumer behaviour studies (Johnson and Stewart 2005; Lefebvre, Hasford, and Wang 2019; Soscia 2007; Watson and Spence 2007), tourism research (Breitsohl and Garrod 2016; Cai, Lu, and Gursoy 2018; Choi and Choi 2019; Hosany 2012; Jiang 2019; Ma et al. 2013), and emerging studies on residents’ emotional states (Ouyang et al. 2017; Zheng et al. 2019a; Zheng et al. 2019b) confirm the merits of cognitive appraisal theories. Unlike dimensional and categorical approaches, appraisal theories account for variations in emotions and offer a rich basis to understand the antecedents and outcomes of emotions.

Recommendation 1. Most psychology, marketing and tourism studies employ categorical and dimensional approaches. Cognitive appraisal theories offer a unifying approach to studying consumer emotions. Method implications dictate the need to establish
conditional guidelines to employ dimensional (amalgamated groupings of positive emotions and negative emotions), categorical (discrete emotional reactions such as love, guilt), or cognitive appraisal approaches in emotion research. The key consideration is the study’s purpose. If specific emotions are not the study’s primary focus, use the dimensional approach. Conversely, designing studies to reveal behavioural responses (e.g., intention to recommend) from a specific emotion (e.g., joy, awe, delight) or emotions of the same valence (e.g., regret and disappointment) should employ a categorical approach. On the other hand, cognitive appraisal offers the best alternative to investigate antecedents of specific emotions.

Appropriateness of adapted self-report emotion measures from psychology


The psychology literature includes human emotion studies employing several perspectives. One research stream identifies universal basic, primary, or fundamental emotions. Izard’s (1977) DES contains 10 subscales representing a person’s frequency experiencing primary emotions (e.g., joy, anger, or disgust). Critics argue DES overemphasises negative emotions, suggesting the scale’s limitation when positive emotions play a central role (see Schoefer and Diamantopoulos 2008). Plutchik’s (1980) psycho-evolutionary theory of emotion consists of eight primary emotion categories and allows
researchers to combine primary emotions to create higher order, more complex emotional states. For example, delight combines joy and surprise.

Although basic emotions offer compelling measures, their use in behavioural research is problematic. Little agreement exists about which emotions are basic, and why they are basic. Emotion theorists find each measurement set proposes different basic emotions (Ortony and Turner 1990). For example, interest and surprise are basic emotions in only Frijda’s (1986), Izard’s (1977), and Tomkins’s (1984) conceptualisations. Arguably, the mechanisms to identify everyday emotions are complex and poorly understood (see Richins 1997). The evidence suggests that validity issues concerning basic emotions theories exist.

Mehrabian and Russell’s (1974) pleasure-arousal-dominance (PAD) captures individuals’ emotional responses to their environment using three independent bipolar dimensions: pleasure /displeasure (P); arousal/non-arousal (A); and dominant/submissive (D). Retailing research commonly employs PAD to assess customers’ in-store emotional experiences (e.g., Donovan et al. 1994; Mazaheri et al. 2014) and studies show the three independent dimensions interact in explaining behaviour (see Massara, Liu, and Melara 2010; Miniero, Rurale, and Addis 2014). However, PAD suffers from two primary weaknesses. First, PAD’s dominance dimension fails to display predictive validity (see Donovan et al. 1994). This validity issue leads researchers to disregard the dominance dimension in conceptualising emotional response (e.g., Sherman, Mathur, and Smith 1997; Walsh et al. 2011). Furthermore, PAD captures global emotion dimensions, and the framework fails to account for discrete emotions.

To address scale reliability and validity issues, Watson, Clark, and Tellegen (1988) develop the “Positive and Negative Affect Schedule (PANAS)”. PANAS’s 20-item self-reporting scale captures positive (PA) and negative (NA) emotions. PA reflects the degree an individual generally feels a zest for life and experiences pleasurable emotions. Conversely,
NA epitomises non-pleasurable engagement such as anxiety and guilt (Watson and Tellegen 1985). Studies employ PANAS to understand satisfaction (e.g., Dubé and Morgan 1998) and post-purchase behaviours (Mooradian and Oliver 1997). Notably, PANAS’s purported subscale independence remains questionable (see Crawford and Henry 2004). PA and NA dimensions appear uncorrelated because measurement errors attenuate negative correlations between dimensions (Green, Goldman, and Salovey 1993). Further, accounting for random and systematic errors reveals a bipolar structure (Barrett and Russell 1998; Carroll et al. 1999).

Some researchers question the relevance of psychology-based emotion scales in consumer (Laros and Steenkamp 2005; Richins 1997; Schoefer and Diamantopoulos 2008) and tourism studies (Hosany and Gilbert 2010; Lee and Kyle 2013). Adapted measures often fail to achieve content validity, leading to erroneous conclusions (Kalafatis, Sarprong, and Sharif 2005; Gilmore and McMullan 2009; Haynes, Richard, and Kubany 1995). Realizing the need to improve measurement validity, influential studies developed context-specific scales assessing consumers’ emotional responses toward ads (viz, Edell and Burke 1987; Holbrook and Batra 1987). Other examples include Richins (1997) “consumption emotion set” (CES) measuring emotions encountered during consumption experiences. Honea and Dahl’s (2005) “promotion affect scale” (PAS) captures consumers’ emotional reactions to sales promotion offers. Schoefer and Diamantopoulos’s (2008) ESRE scale measures emotions during service recovery encounters. Lichte and Plichon (2014) developed a six-dimensional scale to capture emotions experienced in retail outlets.

Hosany and Gilbert (2010) note that existing emotion scales from psychology are context specific and potentially over-represent, omit or under-represent some facets of the tourism experiences. To measure the intensity and diversity of tourists’ emotional responses, Hosany and Gilbert (2010) developed a three-dimensional “Destination Emotion Scale”

Recommendation 2. Researchers improve their studies by first establishing content validity of adapted emotion scales. Depending on the study’s objectives, researchers can develop context specific emotion measures. Tourism researchers tend to borrow and adapt emotion measures from psychology. Since these scales are not developed for tourism research, they unlikely capture the entire domain (type, nature and intensity) of tourism related emotions. Emotional experiences vary from one situation to another and are reliant on the measurement tool. Discipline-relevant scales may not be enough for a wholesale adoption in tourism studies.

Verbal, non-verbal or indirect qualitative emotion measures?

Despite self-report verbal measures’ popularity in emotion research, this method suffers from limitations (Mauss and Robinson 2009). People are unable or unwilling to report true emotions using self-reports (Caruelle et al. 2019), with individuals high in social desirability less willing and/or capable of reporting negative emotional states (Paulhus and Reid 1991). Triangulating verbal and non-verbal measures (psychophysical indices) offer one solution to capture the complexities of emotional states (Lazarus 1991). Prior studies employ psychophysiological techniques to measure emotions, particularly for advertising research (Cacioppo and Petty 1985; Li et al. 2018b). Psychophysiological indices include involuntary measures such as heart rate, blood pressure, facial muscle activity, skin conductance, finger temperature, respiration and eye movement variability (Hadinejad et al. 2019; Kim and Fesenmaier 2015; Li, Scott, and Walters 2015; Parrott and Hertel 1999; Shoval, Schvimer,
and Tamir 2018). For example, facial expressions suggest the presence and intensity of certain emotions (Keltner and Ekman 1996). The Facial Action Coding System (FACS) captures the expression of emotions (Ekman and Friesen 1978). Researchers employ FACS to measure emotional responses to advertisements (see Derbaix 1995).

Electromyography (EMG) also captures measures expression changes. Researchers successfully utilise EMG to measure facial muscle activity (Wang and Minor 2008) and support this method’s validity and reliability (Bolls, Lang, Potter 2001; Hazlett and Hazlett 1999). Another psychophysiological indicator of emotional arousal, Electrodermal activity (EDA), measures the electrical conductance, resistance, impedance, admittance of the skin (Boucsein 2012). Skin conductance remains the most popular type of EDA measurement. Application of EDA is on the rise in consumer studies (see Caruelle et al. 2019 for a comprehensive review). A few studies in tourism (e.g., Kim and Fesenmaier 2015; Shoval et al. 2018) have also successfully employed EDA to measure emotions.

Brain scan technology (viz, functional Magnetic Resonance Imaging or fMRI) also offers exciting opportunities to investigate behaviour and decision-making at the neural level (Casado-Aranda, Sánchez-Fernández, and Montoro-Ríos 2017; Hsu and Cheng 2018; Kenning, Plassmann, and Ahlert 2007; Yoon et al. 2006). Employing fMRI enables researchers to monitor brain activity when consumers are exposed to various marketing stimuli (e.g, brands, or advertisement) during an experiment. Using fMRI, Deppe et al. (2005) investigate the brand information’s influence on brand choice. The fMRI results show subjects exposed to their favourite brands reduce analytic processing and activate brain areas responsible for integrating emotions into the decision-making process. Furthermore, neuroimaging’s application in advertising research shows how the brain processes and stores stimuli. Ambler, Ioannides, and Rose (2000) employ Magnetoencephalography (MEG) to
measure subjects’ responses to affective and cognitive advertisements. Their results suggest emotional content positively affects recall.

Brain imaging technologies, however, have limitations. Researchers typically lack specialist knowledge and experience to use fMRI as a tool, making cross-disciplinary collaborations necessary. Compared to self-reports, designing psychophysiological experiments are more expensive, become time-consuming, involve complex data analysis, and have sample size limitations (Yoon et al. 2006). Interpreting fMRI results also presents a daunting task. The human brain’s one hundred billion brain cells each are equipped with an average of ten thousand connections to other nerve cells, each able to handle one bit of information per second (Nørretranders 1991, p. 143). Brain imaging analysis also opens the door to criticisms about crossing ethical boundaries. Potential flashpoints include mind control and privacy invasion concerns (Wang and Minor 2008).

In addition, indirect qualitative methods provide an effective approach to measure emotions (Lazarus 1991). Yoo, Park, and Maclnnis (1998) demonstrate how ethnography uncovers consumer emotions during shopping experiences. In situ data collection provides an opportunity to observe emotional intensity. The researcher’s role becomes bricoleur—listening to emic accounts and observing nuances to create etic interpretations and developing a gestalt image that explains the subject’s emotional experience (see Denzin and Lincoln 1998). Ethnographic interviews probe deep into unconscious memories offering new insights about emotions. This technique identifies unique emotional responses (e.g., nullification and pride) not typically part of standard emotion typologies (see Yoo et al. 1998).

Qualitative studies are resource-intensive, limiting the researcher’s ability to simultaneously investigate multiple emotions for large samples (Richins 2008). Like other methods, individuals may not feel comfortable discussing sensitive issues, or they may not recall specific events. Methods such as Zaltman metaphor-elicitation (ZMET) may help to
measure and map important aspects underlying consumers’ mental models (see Christensen and Olson 2002; Zaltman 2003; Zaltman and Coulter 1995). Troilo, Gito, and Soscia (2014) successfully utilise ZMET to derive a list of positive emotion theatregoers use to express their feelings toward live performances. Other techniques such as storytelling (Schank 1990) and long interview method (McCracken 1988) help to access information stored unconsciously. Recently, Rahmani, Gnoth, and Mather (2019) introduce Corpus Linguistics (CL) as a novel approach to objectively extract and analyse tourists’ emotional experiences expressed in large volumes of Web 2.0 travel blogs. CL, a branch of linguistics, offers an empirical approach to analyse large collections of ‘real life’ text (Pollach 2012; Rahmani, Gnoth, and Mather 2018).

Some studies attempt to compare and contrast verbal, non-verbal and indirect qualitative emotion measures in tourism. Li et al. (2018a) use self-report and psychophysiological techniques (facial EMG and skin conductance) to measure emotional responses to DMOs advertisements. They show that self-report overestimates the effect of pleasure on advertising effectiveness. In another study, Hadinejad et al. (2019) conclude that combining verbal (self-report), non-verbal (skin conductance, FaceReader™) and indirect qualitative (post hoc interview) provide a more accurate account of ad evoked emotional reactions.

**Recommendation 3.** Both cognitive (verbal self-report and qualitative methods) and psychophysiological measures suffer from a major shortcoming: neither cognitive nor physiological measures explain why a person feels the emotion. If the objective is to assess tourists’ recall of subjective emotional experiences and the resulting behavioural reactions, self-report verbal measures offer the easiest alternative to researchers. For studies investigating emotional responses in greater depth and its complexities, qualitative methods are most appropriate. Ultimately, accessing unconsciously stored emotions require triangulating verbal, non-verbal, and indirect qualitative measures.
Unipolar versus Bipolar Emotion Scales?

Should researchers use unipolar (Likert scales) or bipolar items (semantic differential scales) to measure emotion? Some evidence suggests pleasant and unpleasant emotions are independent (e.g., Diener and Emmons 1984; Cacioppo, Gardner, and Berntson 1997). Measurement errors may mask bipolarity in emotion ratings (see Green et al. 1993). Accounting for random and systematic errors reveals a bipolar structure (Barrett and Russell 1998; Carroll et al. 1999). Unipolar versus bipolar emotion conceptualisation receives little attention and mixed evidence exists in the marketing literature (Babin, Darden, and Babin 1998). Some studies report positive and negative emotions as opposite poles on the same emotional dimension (e.g., Pavelchak, Antil, and Munch 1988; Hui and Bateson 1991). Other scholars question emotion’s bipolarity (e.g., Babin, Darden, and Babin 1998; Jang and Namkung 2009). Prior studies uncover distinct positive and negative emotion dimensions (e.g., Darden and Babin 1994; Mano and Oliver 1993; Oliver 1993).

To resolve bipolarity concerns, a unipolar view provides an alternative. This view distinguishes between ambivalence (joint occurrence of pleasant and unpleasant states) and indifference (occurrence of neither pleasant nor unpleasant states) (Edell and Burke 1987; Westbrook 1987), and enables positive and negative emotions to co-occur (Williams and Aaker 2002). Several consumer behaviour (e.g., Aaker, Drolet, and Griffin 2008; Andrade and Cohen 2007; Lau-Gesk 2005; Olsen, Wilcox, Olsson 2005; Penz and Hogg 2011; Schmalz and Orth 2012) as well as psychology studies (e.g., Berrios, Totterdell, and Kellett 2015; Berrios, Totterdell, and Kellett 2018; Larsen, McGraw, and Cacioppo 2001; Priester and Petty 1996; Schimmack 2001) establish that individuals subjectively experience mixed emotions.
Feeling positive emotions do not preclude the occurrence of negative emotions (Andrade and Cohen 2007; Babin et al. 1998). For example, the tourist is excited about the Paris Catacombs experience, but is frustrated about the long wait to get inside. This situation demonstrates positive and negative emotion co-occurrence (Cacioppo and Bernston 1994). In advertising research, Williams and Aaker (2002) note attitudes towards ads incorporating mixed emotions (e.g., both happiness and sadness) depend on individual’s proclivity to accept duality. Mixed emotions lead to less favourable attitudes for individuals with a lower propensity to accept duality (Anglo Americans, young adults) relative to those with a higher propensity (Asian American, older adults). Mixed emotions also associate with experiences such as white-water rafting (Arnould and Price 1993) and gift exchange (Otnes, Lowrey, and Shrum 1997). Surprisingly, not much is known about mixed emotions effect on outcome variables such as satisfaction and intention to recommend (cf. Williams and Aaker 2002) in the context of tourism.

Recommendation 4. People from certain cultures are more likely to experience mixed emotions. The study settings therefore dictate the measurement of emotions. Bipolar scales obscure differences in emotional responses and do not allow the researcher to capture the co-occurrence of positive and negative emotions. The evidence suggests positive and negative emotions display distinct and asymmetrical effects on behaviour. For empirical studies, theorising emotions under a categorical approach with unipolar rating scales is desirable.

Retrospective vs. In-Process (Online) Emotions?

Studies in tourism mostly use post-visit surveys to measure retrospective emotional states. Retrospective reports are vulnerable to memory reconstructions, increasing the likelihood of recall bias (Chang et al. 2014). Recall-ability declines quickly over time (Weaver and Schwarz 2008). Delays between experiencing and reporting an emotional episode result in random and systematic retrospective biases (Robinson and Clore 2002a;
Random biases occur when respondents selectively retrieve some contextual details. Systematic biases occur because some detail retrieval (e.g., most recent moments of an experience) comes at the expense of missing other aspects of the same experience (Robinson and Clore 2002a).

Depending on the reference time, individuals rely on different strategies to recollect past emotional experience (Robinson and Clore 2002a). Recalling short, discrete time frames (e.g., “How happy were you over the past 30 minutes?”), individuals draw on episodic knowledge (Robinson and Clore 2002b). Individuals recall explicit details from a specific event to judge their prior emotional state (Scollon et al. 2004). Recalling longer and more abstract time frames, individuals use heuristic information or semantic knowledge (e.g., general self-beliefs) (Robinson and Clore 2002a).

Furthermore, emotional experiences fluctuate considerably over time (e.g., Fredrickson and Kahneman 1993). Research establishes emotional responses are dynamic and often change throughout the consumption/tourism experience (e.g., Arnould and Price 1993; Gao and Kerstetter 2018; Kim and Fesenmaier 2015; Lin et al. 2014; Nawijn 2011; Nawijn et al. 2013). In the context of extended service transactions, Dubé and Morgan (1998) find consumers’ emotional states improve gradually due to service provider interactions, suggesting skewed results for global, post-encounter evaluations. Other studies (e.g. Lin et al. 2014; Mitas and Bastiaansen 2018; Mitas et al. 2012; Nawijn et al. 2013) also establish that individuals’ self-reported emotions vary in type and intensity throughout the tourism experience.

Accordingly, memories of emotions often provide inaccurate accounts of online experiences (Levine 1997; Thomas and Diener 1990; Wirtz et al. 2003). Distortions in recall of experiences reflect a positive bias, also known as the “rosy view” phenomenon (Mitchell et al. 1997). The rosy view effect mitigates negative emotions in people’s retrospective
assessments and magnifies positive experiences (Gilbert et al. 1998). One plausible explanation for the positive bias is that, although negative experiences reduce the enjoyment of the moment, these disappointments are fleeting (Mitchell et al. 1997) and people reinterpret their memories in ways consistent with their original expectations (Klaaren, Hodges, and Wilson 1994). In a series of studies, these researchers demonstrate how hedonic evaluations of past experiences are influenced by the most intense (peak) and final (end) moments as opposed to the sum or average of every moment of that experience. The existence of “peak-end” effect is independent of duration (Ariely and Loewenstein 2000) and occurs for both pleasant and unpleasant experiences (Fredrickson and Kahneman 1993; Fredrickson 2000). Recent evidence however suggests tourists in non-hedonic contexts purposely seek and welcome negative emotions (Knobloch et al. 2017; Nawijn and Biran 2019). Negative emotions lead to positive outcomes such as eudaimonic happiness and an identity formation (Nawijn and Biran 2019) and ethical choice formation (Malone, McCabe, and Smith 2014).

To address the biases introduced by retrospection, experience sampling method (ESM) offers a solution for tourism research (Cutler, Doherty, and Carmichael 2018). EMS asks respondents to respond to repeated assessments over the course of time while functioning within their natural settings (Scollon, Kim-Prieto, and Diener 2003). Tourists complete self-reports about their emotional experiences at specific location and time intervals (Shoval, Schvimer, and Tamir 2018). Advances in technology aid the use of smartphones as an ESM tool (Cutler, Doherty, and Carmichael 2018; Kuntsche and Labhart 2013). Several studies demonstrate the convergent validity between aggregated experience sampling data and retrospective emotional reports (e.g., Barrett 1997; Scollon et al. 2004). Furthermore, the “Day Reconstruction Method” (DRM) also aims to minimise or eliminate recall biases in people’s assessments of their feelings and experiences (Kahneman et al. 2004). The DRM
encourages participants to reconstruct a diary consisting of a sequence of episodes. Such episodic re-instantiation procedure aims to evoke one’s contextual experience and attenuate biases commonly associated with retrospective reports (Robinson and Clore 2002a). In addition, diary methods (Bolger, Davis, and Rafaeli 2003) also reduce the likelihood of retrospection by minimising the amount of time elapsed between an experience and the report of this experience. Diary methods allow the researcher to examine reported events and experiences in their natural and spontaneous context (Reis 1994). Studies in tourism have successfully use diary methods to collect data about tourists’ emotional experiences (Cai, McKenna, and Waizenegger 2019; Gao and Kerstetter 2018; Gao et al.2019; Lin et al, 2014; Nawijn et al, 2013).

Recommendation 5. Time obscures and distorts memories. Recalling emotions are difficult for respondents, particularly during abstract time frames. To mitigate retrospective evaluation problems, researchers must adopt strategies to facilitate respondents’ recall of events. Providing cues to respondents improve the likelihood of accurate recall. Where did the event occur? Who was with you at the time? The tourism experience as extended service encounters often become a rollercoaster of emotions. Post-experience global evaluations unlikely capture these discrete emotions. Capturing in-process emotions using techniques such as ESM, DRM and diary methods helps to overcome problems associated with retrospective global reports. Future research needs to establish in-process emotion’s effect on outcome variables such as overall satisfaction, attitudes, and behavioural intention.

Interplay of emotion and cognition in tourist behaviour models

The psychology literature debates whether cognition precedes emotion or emotion precedes cognitive evaluation (see Lazarus 1999). In his pioneering article (Zajonc 1980) and subsequent works (e.g., Zajonc and Markus, 1982), Zajonc posits that emotions have primacy over and are independent of cognitions. On the other hand, the cognition-emotion school of
thought argues cognitions fundamentally determine emotional experiences (e.g., Lazarus 1991). In the marketing literature, Chebat and Michon (2003) and Walsh et al. (2011) tested competing models linking emotions, cognitions and behaviour. Both studies support the proposition that cognitions precede emotions.

Traditionally, influenced by the dominant neo-classical economics paradigm (e.g., Bettman, Luce, and Payne 1998; Sheth, Newman, and Gross 1991), customer satisfaction literature focuses on a cognitive approach suggesting congruence between performance and comparison standards leads to satisfaction—the expectancy disconfirmation model (e.g., Oliver 1980; Wirtz, Mattila, and Tan 2000). Prior studies have challenged cognitive models of customer satisfaction in their ability to predict future behaviour (Bigné, Mattila, and Andreu 2008; Phillips and Baumgartner 2002).

A research stream concludes both cognition and emotional responses to a product or service shape evaluative judgements (e.g, Bigné et al. 2008; Caro and Garcia 2007; Homburg, Koschate, and Hoyer 2006; Ladhari 2007). This research tradition identifies two competing satisfaction models. The first perspective suggests emotions mediate between cognitive evaluations (e.g, perceived product performance) and overall satisfaction (see Bigné et al. 2008; Dubé and Menon 2000; Schoefer 2008). A second approach suggests emotions and cognitions act as independent variables affecting satisfaction judgements (e.g, Mano and Oliver 1993; Oliver 1993).

Recommendation 6. Cognitive models alone offer limited predictive power to account for satisfaction and subsequent behaviours. To overcome this weakness, include both emotions and cognitions when modelling tourist behaviour. A need exists to further understand the interplay and hierarchy of cognitions and emotions in tourism studies. Researchers should theorise and test competing models. Emotional responses could be either an independent variable or a mediator between cognitions and outcome variables (e.g., behavioural intention, and perceived overall image evaluation).

Linear versus configurational theoretical and measurement design

A seventh general consideration is introduced here—the accuracy of measurement from two issue perspectives: (1) does the measure value (score) accurately reflect the salience, valence, and implicit-explicit nature of what is being measured and (2) is the measure accurate in predicting relevant outcomes (e.g., behavioural intention). Measuring, testing, and refining the psychometric properties of emotions respond to the first of the two issue perspectives. The use of symmetric linear models (e.g., multiple regression models (MRA) and structural equation models (SEM) with null hypothesis significance testing (NHST) versus asymmetric configurational models (e.g., fuzzy-set qualitative comparative analysis (fsQCA), see Ragin (2008) for a primer) includes the two streams of attempts to respond to the second issue.

SEM theory and testing are highly useful for proposing and confirm the psychometric properties of scales to measure emotional traits. Separate multiple items measuring each trait should associate with their respective traits and no others. Thus, the distribution of responses for items 1, 2, 3, and 4 to measure trait A should all “load heavily” on the same factor and no others and items 5, 6, 7, and 8 to measure trait B should all load heavily a second factor and not others. SEM testing for trait purity and accuracy for measuring unique traits is the major benefit of this tool/theory.
However, the use of MRA and SEM for proposing and testing variable directional relationships (VDR) using NHST is bad science practice as a number of scholars (Hubbard 2015; Meehl 1978; Woodside 2019; Ziliak and McCloskey 2008) and the American Statistical Association (ASA) advance (Wasserstein and Lazar 2016). Linear symmetric theory construction and data analysis fail to capture the inherent complexities of the impact of emotions on outcome variables. NHST, while still the pervasive logic in emotion research in tourism, is being challenged by the use of odds ratios of predictions of specific interval and point estimates (e.g., Ordanini, Parasuraman, and Rubera 2014; Woodside, Prentice, and Larsen 2015). Woodside (2019) refers to odds ratio predictions of outcomes as “somewhat precise outcome tests” (SPOT).

Applying asymmetric configurational theory and analysis in emotion research supports the suggestion that the configuration of one or a few (i.e., 2 to 5) emotions’ high salience (S), high valence (V) and high attachment (A) is sufficient for accurately predicting a given response (e.g., intention to recommend a specific destination). With E1 representing one emotion trait, equation 1 states this proposition as a Boolean expression:

\[ E_1S \bullet E_1V \bullet E_1A = E_{Total} \leq Y_{Outcome} \] (1)

The mid-level dot (“\(\bullet\)”) in equation 1 represents the logical “AND” condition in Boolean algebra. For example, all measures in fsQCA are transformed from their original values to membership scores ranging from 0.00 to 1.00 via a lazy “S” logistic function. The logical “AND” function calculates a high emotion score (e.g., \(\geq 0.95\)) only if all three parts of an emotion are high—the total score for \(E_{Total}\) is equal to the lowest score in the “fuzzy statement” represented by \(E_1S \bullet E_1V \bullet E_1E\). This operation is helpful for expressing multiple emotional states, for example, considering low, medium, and high states for salience, valence, and attachment, 27 configurations are possible theoretically (3 by 3 by 3 = 27). Only one of these 27 configurations would achieve a high score if “high” is equal or above 0.95;
“moderate” is to 0.50; and a “low” is equal to 0.05. An asymmetric configuration theory of emotion in tourism research might include the algorithm statement that the occurrence of one or a combination of two-to-five emotions achieves a high \( E_{\text{Total}} \) membership score is sufficient for indicating a high outcome score (e.g., a highly positive attitude toward the destination). This configurational design for emotions in tourism research for measurement and predictive theory receives additional elaboration in the “emotionapps model” below.

**Recommendation 7.** Symmetric linear models (multiple regression models and structural equation models) of emotions are prevalent in tourism for theory construction and analysis. Linear models have received criticisms for its failure to capture the complexities of emotions and its impact on outcome variables. To avoid bad science practices, asymmetric configurational modelling (e.g., fsQCA) offers a rich alternative for theory construction and testing in tourism emotion studies. Asymmetric modelling allows researchers to understand how multiple emotions explain and predict the same focal outcome variable.

**The emotionapps model: Salience, valence, and consciousness**

The preceding critical discussion identifies key considerations and offers recommendations to scholars and practitioners considering how best to measure emotional experiences. As a further guide, the emotionapps model provides a roadmap to help researchers choose the right emotion measures (see Figure 1). The emotionapps model includes the proposal that all combinations (configurations) of high/low salience, negative/positive valence, and conscious/unconscious are possible and do occur depending in part on contextual frames. The emotionapps model improves upon Bargh’s (2002) perspective that consciousness/unconsciousness is a continuum—the individual often does have some (but incomplete) ability to interpret some aspects of her/his unconsciousness. Viewing consciousness/unconsciousness as a configuration where upon the individual combines conditional bits of
information called up from short- and long-term memory and contextual cues improves on Bargh’s (2002) continuum perspective.

Thus, Figure 1 includes the mid-location “I” in recognition that moderate levels of salience, valence, and consciousness sometimes (possibly frequently) occur. Consequently, a theoretical “property space” (Lazarsfeld 1937) analysis of the emotionapps model indicates that nine spaces have relevancy for developing emotion metrics—covering low, medium, and high levels of each of the three dimensions. Property space analysis depicts a truth-table view of all possibilities involving combinations of conditions or factors. To use emotionapps model, the researcher needs a holistic understanding of the study context. Who is being studied? What does the researcher know about the study group? What circumstances or experimental conditions exist? The Emotionapps model’s key dimensions are salience, valence, and consciousness.

Figure 1 here.

Emotional salience refers to an assessment of the significance of a stimulus to the individual (Berenbaum, Boden, and Baker 2009). Evaluative judgments high (low) in salience lead to positive (negative) emotions. For researchers interested in studying causes of emotions (salience), appraisal theories offer a solution. Appraisal theories characterise emotion as a mental state arising from subjective evaluations and interpretations of events on a number of dimensions (Roseman et al. 1990). For example, appraising goal congruence is an assessment of whether a situation is conducive to goal fulfilment (Hosany 2012). A close link exists between people’s goals and the emotions they experience (Carver and Scheier 1990).
Goal congruent situations lead frequently to positive emotions and goal-incongruent situations generate negative emotions. Salience also enables researchers to study causes of negative emotions (such as regret and disappointment). Three key appraisal dimensions are relevant to understand negative emotions: fairness, coping potential and agency. The appraisal of fairness refers to the extent one perceives an event as appropriate and fair (Frijda 1986). Coping potential reflects an individual’s evaluation of the potential to cope with a situation to attain a desired outcome or avoid an undesired one (Roseman et al. 1990). Finally, appraising involves the attribution of cause (oneself, someone else or circumstance) to an outcome (Ortony, Clore, and Collins 1988).

Valence is the subjective feeling of pleasantness or unpleasantness (Barrett 1998). Affective experiences of low valence are laden with negative emotions and high valence indicates positive emotions. A long-standing debate exists among emotion theorists about whether valence is irreducible (e.g., Russell and Carroll 1999) or whether positivity and negativity are separable in experience (e.g., Cacioppo and Berntson 1994). Can people feel happy and sad at the same time? Consumer studies provide support for the occurrence of mixed emotions (e.g., Arnould and Price 1993; Otnes et al. 1997). Mixed emotions have meaningful consequences (Larsen and McGraw 2011).

For example, evidence indicates advertisements eliciting mixed emotions influence their effectiveness (Williams and Aaker 2002). Some people may experience mixed emotions more often than others; Asians have a tendency to experience more mixed emotions than Westerners (e.g., Bagozzi et al. 1999; Schimmack 2009; Williams and Aaker 2002). For a more complete understanding of emotional experience in bittersweet situations containing both pleasant and unpleasant aspects (Larsen et al. 2001), requires conceptualising positive and negative emotions as separable (Larsen and McGraw, 2011). As a result, theorising emotions under a categorical approach with unipolar measurement scales is desirable.
Emotional consciousness refers to the extent a person is aware of his/her emotions (Roberts 2009; Thagard and Aubie 2008). Verbal self-report measures are sufficient when awareness is high (position C) and emotions are captured in-situ. But, do people always know how they feel? Individuals are not always conscious of their emotions and may not understand their own feelings (Winkielman, Berridge, and Wilbarger 2005; Zaltman 2003). Several theorists have emphasised the importance of having access to one’s feelings, the ability to discriminate among these feelings, and being able to label one’s feelings (e.g., Berenbaum, Boden, and Baker 2009). Unconsciously, the memory stores life experiences, but an individual may not know why they feel a specific emotion (Bargh 2002; Wilson 2002).

Panksep and Biven (2012) further conclude basic emotions do not emerge from the cerebral cortex associated with complex thought; instead, these feelings generate from deep, ancient brain structures. The amygdalae and hypothalamus are thoughts to contain emotions uses non-verbal, involuntary measures. Panksep and Biven (2012) findings follow the proposition that most human thinking occurs unconsciously (Wegner 2002; Zaltman 2003). In low emotional consciousness situations (position D), psychophysiological techniques offer the best approach to measure emotions. With technological advances, psychophysiological techniques offer a popular, new data source to measure emotions. Psychophysiological methods do not require cognitive processing and provide a more comprehensive, unbiased account of an individual’s reaction to a stimulus.

Conclusions

A rich body of research establishes emotion’s relevance in tourism. Emotions play an important function in defining tourism experiences and influencing tourist evaluations. Prior studies predominantly adapt existing self-report emotion scales from psychology. Concerns are surfacing about the applicability, reliability, and validity of psychological emotion measures in tourism. Since these adapted self-report scales were not developed
purposely for tourism, they underrepresent tourism related emotions suggesting important shortcomings. Table 1 summarises the recommendations for methodological and theoretical issues that researchers may encounter.

Table 1 here.

The complexity of usefully measuring emotions cannot be overstated in tourism research. The present article synthesise the cross-disciplinary emotion literature and addresses methods concerns as well as provides guidance to select highly useful-for-the-context (HUFTC) measures. Previous studies suggest a magic bullet does not exist for emotion studies. Researchers considering psychology literature measurement scales should proceed with caution. Either the scale’s content validity needs to be established or new scales need development and validation to address specific study objectives. Verbal self-reports remain the most popular method to capture tourists’ emotional experiences, but they suffer from some serious limitations. Psychophysiological measures, on the other hand, do not require cognitive processing and tap into the richness and complexity of the emotional response. Asymmetric modelling allows researchers to understand how multiple emotions predict the same outcome variable. Triangulating several methods (verbal, non-verbal, and qualitative emotion measures) helps to capture the complexities of emotional states.

Measuring in-process emotions mitigate random and systematic retrospective biases associated with post-experience surveys. If in situ data collection is not possible, researchers need to develop strategies to help tourists recall their experiences. Providing cues to trigger memories during specific time frames offers a reasonable alternative. Emotion-triggering experiences elicit both positive and negative emotions. Unipolar scales best capture the co-occurrence of positive and negative emotions.
The study here extends the body of literature, highlighting several issues in measuring emotions, and formulates a number of recommendations to aid future research. Emotion’s roots are ecological and knowledge about a person’s environment helps explain why people feel the way they do (Mathur and Moschis 2005). Such limitation also offers opportunities for new directions in understanding how emotions affect tourist behaviour. Emotion suppression is likely to occur, and tourists’ feelings are stored unconsciously (see Bargh 2002). These unconscious memories influence behaviour. Future research needs to focus on providing additional tools useful for revealing memories stored unconsciously to gain a more comprehensive insight on tourists’ emotions. Finally, this study, synthesising multiple literature streams, focuses on the methodological and theoretical issues in tourism emotion research. Future studies may provide a complete review of the wider role emotions play in tourism.
Figure 1
Emotionapps: Salience, Valence and Consciousness Perspectives
<table>
<thead>
<tr>
<th>Methodological and Theoretical Issues</th>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Are summary dimensions of emotion appropriate?</td>
<td>If specific emotions are not the study’s primary focus, use the dimensional approach. Studies designed to reveal behavioural responses from a specific emotion or emotions of the same valence should employ a categorical approach. For investigations into antecedents of specific emotions, cognitive appraisal offers the best alternative.</td>
</tr>
<tr>
<td>2. Are adapted self-report emotion measures from psychology appropriate?</td>
<td>Emotion taxonomies from psychology are not conceived to categorise and measure tourists’ emotional experiences per se. Adapting emotion scales from psychology fail to achieve content validity and leads to erroneous conclusions. Researchers must first establish content validity of borrowed emotion scales. Depending on the study objectives, researchers can develop context specific emotion measures.</td>
</tr>
<tr>
<td>3. Should verbal, non-verbal, or indirect qualitative emotion measures be employed?</td>
<td>Self-report verbal measures offer the easiest alternative to researchers assessing tourists’ recall of subjective emotional experience and resulting outcomes. Investigating emotional responses in greater depth and its complexities, require qualitative methods. Ultimately, accessing unconsciously stored emotions require triangulating verbal, non-verbal, and indirect qualitative measures.</td>
</tr>
<tr>
<td>4. Unipolar or bipolar emotion scale?</td>
<td>People from certain cultures are more likely to experience mixed emotions. Bipolar scales obscure differences in emotional responses and do not allow the researcher to capture the co-occurrence of positive and negative emotions. The evidence suggests positive and negative emotions display distinct and asymmetrical effects on behaviour. For empirical studies, theorising emotions under a categorical approach with unipolar rating scales is desirable.</td>
</tr>
<tr>
<td>5. Capturing retrospective or in-process emotions?</td>
<td>Recalling emotions are difficult for respondents, particularly during abstract time frames. To mitigate retrospective evaluation problems, researchers must adopt strategies to facilitate respondents’ recall of events. Providing cues to respondents improve the likelihood of accurate recall. The tourism experience as extended service encounters often become a rollercoaster of emotions. Post-experience global evaluations unlikely capture these discrete emotions. Capturing in-process emotions using techniques such as ESM, DRM and diary methods allow full emotional accountability.</td>
</tr>
<tr>
<td>6. How should the interplay of emotion and cognition be used?</td>
<td>Cognitive models offer a limited ability to account for satisfaction evaluations and subsequent behaviours. To overcome this weakness, include both emotions and cognitions when modelling tourist</td>
</tr>
</tbody>
</table>
in tourist behaviour models? Emotion could be either an independent variable or a mediator between cognitions and outcome variables. Researchers should theorise and test competing models.

7. Linear versus configurational theoretical and measurement design? Symmetric linear models (multiple regression models and structural equation models) of emotions are prevalent in tourism for theory construction and analysis. Linear models have received criticisms for its failure to capture the complexities of emotions and its impact on outcome variables. To avoid bad science practices, asymmetric configural modelling (e.g. fsQCA) offers a rich alternative for theory construction and testing in tourism emotion studies. Asymmetric modelling allows researchers to understand how multiple emotions explain and predict the same focal outcome variable.
References


