Emotional Response Coherence and Interoceptive Awareness: Development and Validation of a Novel Assessment Method

By

Luma Muhtadie

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Psychology in the Graduate Division of the University of California, Berkeley

Committee in charge:

Professor Robert W. Levenson, Chair
Professor Iris Mauss
Professor Ronald E. Dahl

Summer 2018
Abstract

Emotional Response Coherence and Interoceptive Awareness: Development and Validation of a Novel Assessment Method

by

Luma Muhtadie

Doctor of Philosophy in Psychology

University of California, Berkeley

Professor Robert W. Levenson, Chair

Interoceptive awareness (IA), the conscious perception of signals originating in the body, is a fundamental component of our subjective experience of emotion and may be its proximate cause. IA is integral to attention, motivation, emotion regulation, and decision-making—processes that are essential to our survival, sense of agency, and wellbeing. A clear understanding of individual differences in IA is currently hampered by the limitations of prevailing assessments—namely, self-report questionnaires and heartbeat perception tasks—which have questionable reliability and validity, fail to capture the full spectrum of individual variability, and disregard the emotional contexts in which interoceptive processes naturally unfold. This dissertation proposes a novel method for assessing IA that capitalizes on emotional response coherence. Specifically, the method assesses variability in the extent to which physiology and subjective experience track together within individuals while they are experiencing strong emotions. Theoretical and empirical rationales for considering the “Coherence Task” to be a proxy measure of IA are elucidated and its psychometric properties are systematically evaluated.

Fifty-six men and women aged 18 to 50 completed the coherence task on two occasions spaced one week apart. While watching evocative film montages that captured a range of emotions, subjects provided momentary ratings of their subjective experience on valence and arousal dimensions (2 separate trials per session) and their physiology was continuously recorded. Cross-correlation coefficients of the coherence between subjective ratings and heart period (“Coherence Scores”) were then computed for each individual. Coherence Scores derived from valence-based ratings of subjective experience and heart period demonstrated significant 1-week test-retest reliability (i.e., temporal stability); were positively associated with self-reported body awareness (i.e., convergent validity); and were negatively associated with a composite measure of distress and positively associated with empathy (i.e., predictive validity). Moreover, these findings showed specificity for the coherence between subjective experience and visceral over somatic signals (i.e., for interoceptive over proprioceptive awareness; discriminant validity).

The Coherence Task shows early promise as an empirically grounded assessment of individual differences in IA. This task would also enable us to evaluate the efficacy of interventions that target interoceptive awareness for health and wellbeing (e.g., mindfulness meditation).
For Michelle McKenzie, whose friendship is a living, breathing teacher of acceptance and love.
With endless gratitude to my mentor, Bob Levenson. For steering me toward meaningful questions (“step away from the literature and watch the behavior”), for teaching me how to play on and with a team, for inviting me to truly consider the role of emotion—not just as a scientist, but as a human being, and for helping me get my groove back.
Introduction

In the broadest terms, interoception refers to the processing of sensory input originating in the body. Interoception is a fundamental component of our subjective experience of emotion (Lakoff, 1987; Levenson, 1999) and may even be its proximal cause (Craig, 2002, 2009; Damasio, 1999; James, 1884). Interoception is also centrally involved in homeostasis, attention, stimulus-response learning, motivated behavior, and decision-making—processes that are essential to our survival, sense of agency, and psychological wellbeing (Craig, 2010; Domschke, Stevens, Pfleiderer, & Gerlach, 2010; Dunn, Galton, et al., 2010; Farb et al., 2015; Kever, Pollatos, Vermeulen, & Grynberg, 2015; Park & Tallon-Baudry, 2014; Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003). Empirical evidence demonstrates that individuals vary considerably in the extent to which they are consciously aware of the information emanating from their bodies from moment to moment (Critchley, Wiens, Rotshtein, Öman, & Dolan, 2004; Katkin, 1985; Pollatos & Schandry, 2004; Schandry, 1981; Wiens, Mezzacappa, & Katkin, 2000). This conscious perception of subtle internal body sensations, termed ‘interoceptive awareness,’ is assumed to be trait-like and thus stable over time (Garfinkel & Critchley, 2013), though this has not been demonstrated empirically. Given the crucial role of interoception in survival, subjective experience, behavior, and psychological wellbeing, a clear understanding of individual differences in interoceptive awareness, and how they relate to functional outcomes, would make an important contribution to multiple fields, including psychology, neuroscience, and behavioral medicine. Unfortunately, clarity is hampered by the limitations of our prevailing assessments—namely, self-report questionnaires and heartbeat perception tasks (to be reviewed). These approaches fail to capture the full spectrum of individual variability in interoceptive awareness and have poor ecological validity insofar as they disregard emotion—the context in which interoceptive processes and our awareness of them naturally unfold. The present study proposed a novel method for assessing interoceptive awareness that draws on emotional response coherence. This “Coherence Task” capitalizes on individual differences in the degree to which subjective emotional experience and physiology track together within a person while he or she is experiencing strong emotions. In this dissertation, I present the theoretical rationale and core features of the Coherence Task and explain why it can be considered a proxy measure of interoceptive awareness. I then evaluate its psychometric properties, including temporal stability and convergent, predictive, and discriminant validity.

Emotional Response Coherence and its Relation to Interoceptive Awareness

Emotional response coherence refers to the notion that emotions organize and synchronize different response systems in such a way that when individuals are in the throes of an emotion, their subjective, behavioral, and physiological responses track together more closely than they do when individuals are at rest (Darwin, 1872; Ekman, 1992; Lazarus, 1991; Levenson, 1994; Tomkins, 1962). To understand the connection between emotional response coherence and interoceptive awareness that I propose here, it will be helpful to first consider the sequence of events involved when an emotion unfolds. A stimulus in the environment becomes salient and an elicitor of emotion when it signals prototypical problems, challenges, or rewards, such as novelty, threat/safety, or pleasantness/unpleasantness. When such saliency is signaled, it engenders within the individual an “emotional response package” (Levenson, 2003) comprising a
variety of components—from subtle and internal autonomic adjustments, including changes in heart rate, breathing, and sweat gland activity, to large-scale observable phenomena, such as facial expressions, vocal utterances, physical gestures, and approach/avoidance behaviors. Interoceptive awareness describes the extent to which a person is consciously of and accurately perceives components of the first type (i.e., autonomic signals from the body). Interoceptive awareness, which is strongly influenced by mental processes such as attention, memory, beliefs, and appraisals (Mehling et al., 2009) and varies across individuals as a function of multiple factors including biology, learning, and body-awareness training (Farb et al., 2015; Singer, Critchley, & Preuschoff, 2009) ultimately gives rise to subjective experience.

The idea that interoceptive awareness produces downstream subjective emotional experience is not new and can be traced back as far as the James–Lange theory of emotion (James, 1884; Lange, 1885). This theory postulates that the perception of afferent information from the body, in conjunction with behavior in particular situations, forms specific emotional reactions, such as anxiety or anger. The James–Lange view has since been extended by contemporary emotion research. For example, Levenson (1999) proposed that the “subjective experience of a given emotion derives largely from the sensations that are generated by the activation of the associated response package” (p. 496). Damasio (1994) posited that emotions function to bring autonomic bodily processes into awareness, providing “somatic markers” that guide decisions and actions. Accordingly, the subjective experience of emotion relies on brain regions that both represent and regulate our continuously changing internal states.

Recent neuroscientific evidence has begun to elucidate the anatomical details of this model by identifying specific regions—primarily the insular cortex and anterior cingulate cortex—that are involved in representing visceral signals and integrating them with higher-order processes, such as goal monitoring, behavioral control, and predicting the outcomes of possible actions (Craig, 2004, 2010; Critchley, 2004; Singer et al., 2009). It is not coincidental that in addition to mapping internal viscerosensory states, the insula and anterior cingulate cortex are also involved in processing cardiovascular arousal and subjective feelings (Craig, 2002, 2003, 2004; Critchley, Corfield, Chandler, Mathias, & Dolan, 2000; Critchley, Mathias, & Dolan, 2001; Critchley et al., 2004; Damasio, 1999; Damasio et al., 2000; Pollatos, Gramann, & Schandry, 2007; Pollatos, Kirsch, & Schandry, 2005a; Pollatos, Schandry, Auer, & Kaufmann, 2007; Zaki, Davis, & Ochsner, 2012).

From the peripheralist view outlined above, it follows that a measure that assesses how closely subjective experience and internal body signals track together in the presence of salient and evocative stimuli—that is, a measure of the coherence between subjective experience and objective physiology during an emotional episode—would ipso facto constitute an index of interoceptive awareness. This thesis has two corollaries. First, because interoceptive awareness is assumed to be a relatively stable trait-like quality of individuals, its proposed proxy (i.e., the coherence between subjective experience and autonomic physiology) should also be reliable within individuals over time. Second, because interoceptive awareness is assumed to produce downstream subjective experience, variability among people in interoceptive awareness would be expected to manifest in distinct social-emotional tendencies. Indeed, individual differences in interoceptive awareness have been consistently associated with several outcomes, including emotional awareness (e.g., alexithymia), psychological adjustment (e.g., anxiety, depression, perceived loneliness, wellbeing), and socio-emotional sensitivity (i.e., empathy). If my proposal to operationalize interoceptive awareness in terms of the coherence between subjective experience and autonomic physiology is sound, I would therefore expect emotional response
coherence to be related in similar ways to these outcomes.

In the present investigation, I seek to test these hypotheses. I begin with a brief review of the prevailing methodologies we use to assess interoceptive awareness and highlight their limitations. I review the emotional response coherence literature and distinguish between two distinct paradigms that have been used to response coherence—the between-subjects approach and the within-subjects approach—and explain why only the latter is appropriate for the present goal of measuring individual differences in interoceptive awareness. I then elaborate on the specific features of a within-subjects measure of emotional response coherence that would render it a proxy measure of interoceptive awareness. Next, I review several functional outcomes related to emotional awareness, psychological adjustment, and socio-emotional sensitivity that have been empirically associated with interoceptive awareness. Finally, I present a series of tests to examine whether a Coherence Task that measures the extent to which autonomic physiology and subjective experience track together within individuals during an emotional episode offers a psychometrically sound way to assess interoceptive awareness. Specifically, I evaluate the Coherence Task for (a) temporal stability, (b) convergent validity, i.e., association with self-reported interoceptive awareness, (c) predictive validity, i.e., association with outcomes that have been associated with interoceptive awareness in the literature, and (d) discriminant validity, i.e., specificity for the coherence between subjective experience and visceral physiology, which taps interoceptive processes (vs. subjective experience and somatic physiology, which taps proprioceptive processes).

Limitations of Prevailing Measures of Interoceptive Awareness

The most commonly used measures for assessing interoceptive awareness are self-report questionnaires and heartbeat perception tasks and both are beset with problems. Perhaps most concerning in regard to self-report questionnaires, typically termed measures of ‘body awareness’, is that there is no widely accepted unifying measurement definition of this construct and therefore no gold standard for assessing the criterion validity of these measures (Mehling et al., 2009). In addition, most validated self-report measures are dominated by an outdated conceptualization of body awareness, largely rooted in a long history of research within psychosomatic medicine and psychopathology, which conflates body awareness with somatosensory amplification and assumes this leads to hypochondriasis and other maladaptive outcomes, such as anxiety and psychosomatic disorders (Barsky, 1992; Clark et al., 1997; De Berardis et al., 2007; Gregor & Zvolensky, 2008; Ludewig et al., 2005; Olatunji, Deacon, Abramowitz, & Valentiner, 2007). In fact, many self-report measures of this sort (e.g., The Body Sensations Questionnaire, Chambliss, Caputo, Bright, & Gallagher, 1984; the Autonomic Perception Questionnaire, Mandler, Mandler, & Uviller, 1958; and the Somatic Perception Questionnaire, Stern & Higgins, 1969) were exclusively developed to assess for anxiety, and phrase items in a way that presupposes sensations will be appraised as uncomfortable or threatening. For example, the Body Sensations Questionnaire (Chambliss et al., 1984) was developed from interviews with agoraphobic patients and lists sensations associated with autonomic arousal that patients reported experiencing during exposure to distressing phobic situations (e.g., “heart palpitations”, “feeling short of breath”, “feeling disconnected from your

---

1 Whereas interoceptive awareness involves the perception of sensations from inside the body, including physical sensations related to internal organ function (e.g., heartbeat, respiration, gastric activity), proprioceptive awareness involves the perception of joint angles, muscle tensions, movement, posture, and balance.
Heartbeat perception tasks are a vast improvement over self-report measures in that they assess awareness of bodily signals objectively, but they are riddled with problems of their own. Heartbeat perception tasks can be classified into two broad types: heartbeat counting tasks and heartbeat discrimination tasks. In heartbeat counting tasks, subjects are asked to sense and count their heartbeats during brief, fixed periods of time of varying length (e.g., Herbert, Ulbrich, & Schandry, 2007). In heartbeat discrimination tasks, a visual or auditory signal triggered by the subject’s own heartbeat is played with either a minimal or prolonged delay, and subjects are asked on each trial to judge whether this exteroceptive signal is synchronous or asynchronous with their own heartbeat (e.g., Jones, 1994; Whitehead, Drescher, & Heiman, 1977; Wiens & Palmer, 2001; Yates, Jones, Marie, & Hogben, 1985). In both versions, subjects are classified as either ‘good heartbeat perceivers’ or ‘poor heartbeat perceivers,’ depending on whether their score falls above or below a predetermined greater-than-chance cutoff. One major problem with this approach is that the resulting dichotomous scores treat all individuals above or below the cutoff as equivalent to each other, muting the spectrum of individual differences within each category (i.e., ‘good’ vs. ‘poor’ perceivers) and exaggerating differences between individuals whose scores are close to each other but fall on different sides of the cutoff. Moreover,
categorizing a continuous predictor, irrespective of how this is done, always results in the loss of statistical power (Aiken & West, 1991).

Another problem with heartbeat perception tasks is that individuals may unwittingly cheat on them by becoming inadvertently aware of their peripheral pulse during the experiment (e.g., if their finger is resting on the arm of a chair), and this has generally not been controlled for in study procedures. Most problematic, however, is the fact that subjects typically perform no better than chance on heartbeat perception tasks. Across all studies—regardless of the particular method used, the sample size and characteristics, or the research question—the frequency of ‘good detectors’ rarely exceeds 40 percent (Brener, Liu, & Ring, 1993; Eichler & Katkin, 1994; Jones, O'Leary, & Pipkin, 1984; Khalsa et al., 2008; Knapp-Kline & Kline, 2005; Ring & Brener, 1992; Rouse, Jones, & Jones, 1988; Schneider, Ring, & Katkin, 1998; Whitehead et al., 1977; Wiens & Palmer, 2001; Yates et al., 1985). In fact, participants frequently report that they were guessing during the task (e.g., Critchley et al., 2004; Wiens, 2005) and investigators often wind up excluding large numbers of ‘poor perceivers’ to create equal numbers in the two groups (Pollatos, Herbert, Matthias, & Schandry, 2007; Pollatos, Kirsch, & Schandry, 2005b), yielding study samples that do not represent the population.

A major clue to the low rates of better-than-chance performance obtained on heartbeat perception tasks lies in the fact that they are administered in the absence of emotional stimuli. Heartbeat perception tasks were designed to focus attention on body sensations, but they do so in an artificial and reductionist way. Whereas in real-world contexts, visceral signals emerge and become accessible to awareness in the context of and in relation to salient stimuli, heartbeat perception tasks are typically performed at rest and ask participants to compare their heartbeats to lights and tones—stimuli that are of little value to individuals and thus unlikely to elicit strong emotional and autonomic responses (Bechara & Naqvi, 2004). In other words, heartbeat perception tasks fail to capture interoceptive processes in the way they are experienced and perceived in the real world.

In sum, the reliability and validity of self-report measures and heartbeat perception tasks remains controversial. None has emerged as the “gold standard” for measuring interoceptive awareness (Khalsa, Rudrauf, Sandesara, Olshansky, & Tranel, 2009; Knapp-Kline & Kline, 2005; Mehling et al., 2009; Pennebaker & Hoover, 1984), and estimates of interoceptive awareness derived from these various approaches are virtually unrelated to each other (Critchley, 2004; Pennebaker & Hoover, 1984; Whitehead et al., 1977). These measures have poor ecological validity, asking participants to self-rate the accuracy with which they perceive bodily sensations or to monitor and detect specific visceral signals in emotionally sterile contexts. The association between self-reported body awareness or performance on heartbeat perception tasks with interoceptive awareness during emotional episodes has yet to be established.

**Measuring Response Coherence**

Despite strong theoretical arguments in favor of the response coherence postulate, empirical support has been rather inconsistent (Barrett, 2006), with some studies finding positive associations among emotion response components (e.g., Dan-Glauser & Gross, 2013; Mauss, Levenson, McCarter, Wilhelm, & Gross, 2005; Sze, Gyurak, Yuan, & Levenson, 2010); others finding weak (e.g., Bonanno & Keltner, 2004; Reisenzein, 2000; Smith, Hubbard, & Laurenceau, 2011) or no associations (e.g., Edelmann & Baker, 2002; Hessler & Katz, 2007; Jakobs, Mansteaed, & Fischer, 2001; Mauss, Wilhelm, & Gross, 2004; Reisenzein, Bördgen, Holtbernd,
and still others finding negative associations (e.g., Buck, 1977; Hubbard et al., 2004; Lacey, 1967; Lang, 1988). This has led several authors to conclude that the coherence postulate is unfounded (Barrett, 2006; Bradley & Lang, 2000; Reisenzein, 2000).

On close examination of the coherence literature, however, two distinct paradigms emerge: the between-subjects paradigm and the within-subjects paradigm (Buck, 1980; Mauss et al., 2005; Sze et al., 2010). In the between-subjects paradigm, a participant who reports greater than average emotional experience would be expected to show greater than average behavioral and physiological responses. Studies taking the between-subjects approach have yielded disparate findings, with some obtaining positive associations—particularly between experience and behavior (e.g., Fischer & Roseman, 2007; Zeelenberg & Pieters, 2004), and others obtaining weak or no associations—particularly between experience and physiology (e.g., Borkovec, Stone, O’Brien, & Kaloupek, 1974; Grossman, Wilhelm, Kawachi, & Sparrow, 2001; Mauss et al., 2004; Mauss, Wilhelm, & Gross, 2003; Weinstein, Averill, Opton, & Lazarus, 1968).

Conceptually, it has been argued that a between-subjects approach is irrelevant to the question of how tightly responses are linked within an individual over time (Buck, 1980; Cacioppo et al., 1992; Lacey, 1967; Stemmler, 1992). Moreover, in this approach sources of between-individual variance are likely to eclipse potential associations among response systems within an individual, making the latter very difficult to detect (Lazarus, Speisman, & Mordkoff, 1963; Pennebaker, 1982; Reisenzein, 2000; Rosenberg & Ekman, 1994; Ruch, 1995).

The alternative within-subjects approach examines the extent to which responses are coordinated within individuals during an emotional episode. Using this approach, we would expect to see greater physiological and behavioral responding during periods when an individual reports experiencing more intense emotion than during periods when the same individual reports experiencing less intense emotion. Only three studies to date have taken a within-subjects approach to assessing response coherence during emotional episodes (Dan-Glauser & Gross, 2013; Mauss et al., 2005; Sze et al., 2010) and these studies obtained positive correlations between subjective experience and physiology that had eluded between-subjects studies. Still, effect sizes remained modest due to high individual variability in emotional response coherence, which one study found to be attributable to differences in interoceptive awareness, albeit indirectly (Sze et al., 2010). Specifically, this study compared coherence across three groups of individuals with varying levels of interoceptive awareness training—experienced vipassana meditators (highest level), experienced dancers (intermediate level), and controls with no former meditation or dance experience (lowest level; Sze et al., 2010). The researchers measured within-individual coherence between momentary ratings of subjective experience (valence-based continuum from very negative to very positive) and continuous heart period in the three groups of subjects (i.e., meditators, dancers, controls) while they viewed a series of emotionally evocative film clips. Results showed a linear pattern of emotional response coherence, with meditators having the highest mean level of coherence, dancers having an intermediate level, and controls having the lowest level. Although this finding suggests that coherence performance across the three groups reflected differences in their interoceptive awareness, the evidence remained indirect insofar as it rested on an assumption about interoceptive ability across the three groups based on their backgrounds in body awareness training.
Key Features of a Response Coherence Measure Constituting a Proxy for Interoceptive Awareness

Clearly, there is a strong argument for using a within-subjects approach to assess emotional response coherence. But for a within-subjects measure of coherence to serve as a suitable proxy for interoceptive awareness, it would need to include several additional features that I enumerate here.

First, the measure would need to assess coherence between two specific channels: autonomic physiology (i.e., the objective body signals that form the substrate of interoceptive awareness) and subjective experience (i.e., the emergent property of this awareness).

Second, within-subject coherence between these two channels would need to be assessed while individuals are experiencing strong emotions. As outlined earlier, functionalist accounts of emotion that posit coherence predict close coordination among response systems during emotional episodes (Davidson, 1992; Levenson, 1994) and less coordination when individuals are at rest (Lacey, 1967; Lazarus et al., 1963)—a prediction that has been borne out empirically (Mauss et al., 2005). Converging with this view, studies of interoceptive awareness suggest that there is an inherent limitation in individuals’ ability to detect visceral sensations while they are at rest (Jones & Hollandsworth, 1981; Karsdorp, Kindt, Rietveld, Everaerd, & Mulder, 2009; Khalsa et al., 2009; Schandry, Bestler, & Montoya, 1993), which can be overcome under conditions of arousal, as occurs in the context of emotion.

Third, the physiological index whose coherence with subjective experience is being assessed should be a prominent source of visceral sensation that is accessible to conscious awareness. So which physiological indices might be considered good candidates? If the metaphors we use to describe our emotions reflect their underlying physiology, as past linguistic and psychophysiological research suggests (Heelas, 1996; Lakoff, 1987; Marchitelli & Levenson, 1992; Pennebaker, 1982; Pérez, 2008), then the wealth of cross-cultural emotion metaphors conceptualizing the heart as the central locus of feelings (e.g., “broken heart”, “heart-throb”, “from the bottom of my heart”) point to this organ as one such prominent source. Moreover, the heart is affected by the intensity of both positive and negative emotions (Bradley & Lang, 2000)—and perturbations from rest in both directions contribute to interoceptive processes. Yet another rationale for focusing on the heart is that the vast majority of existing studies on interoceptive awareness have focused on cardiac parameters (i.e., heart rate and heart period; Kindermann & Werner, 2014; Schandry et al., 1993).

Fourth, the measure should be able to capture the temporal resolution, timing (i.e., onset, duration, offset), and coordination of the subjective and physiological channels, and this is only possible if both channels are assessed continuously. Although obtaining continuous measurement is straightforward when it comes to physiology (i.e., physiological signals are typically measured this way), the issue is somewhat more complicated when it comes to subjective experience. Because obtaining moment-to-moment self-ratings of emotion risks impeding the natural trajectory of emotion (Gottman & Levenson, 1985; Rosenberg & Ekman, 1994), researchers have tended to rely on retrospective and aggregated ratings, which are prone to memory and self-presentational biases (Barrett, 1997; Kahneman, 2000). To address this issue, Levenson and colleagues developed a rating dial methodology for obtaining continuous ratings of subjective experience when studying emotion during dynamic marital interactions (Gottman & Levenson, 1985; Levenson & Gottman, 1983; Ruef & Levenson, 2007). Their “affect rating dial” has since been used to gather continuous ratings of subjective experience in studies examining emotional
response coherence (Dan-Glauser & Gross, 2013; Mauss et al., 2005; Sze et al., 2010), and there is evidence that the method does not alter the natural course of affective responding (Mauss et al., 2005).

The fifth feature pertains to the way the relationship between response components (i.e., between measures of physiology and subjective experience) is characterized. Butler and colleagues (2014) have offered a persuasive argument for using time-lagged cross-correlations to calculate the coherence of time series emotion data. This approach takes into account both the unique time courses of the two response channels and the inherent non-stationarity of this type of data (i.e., the mean and variance of each time series will vary over time with fluctuations in emotional arousal). Non-stationarities are often removed to conduct hypothesis tests that assume stable distribution properties, which has the effect of washing out emotion-based changes in the distribution of data over time—the information that is of principal interest (Gottman, 1981). Computing lagged cross-correlations between subjective experience and physiology for each individual picks up between-variable associations due to shared mean, slope, or variance changes and due to shared oscillations or momentary fluctuations, providing a sensitive within-subject index of emotional response coherence that can be used for between-person hypothesis testing with various outcomes. Because hypothesis testing is done only at the between-person level—i.e., associations between Coherence Scores (cross-correlations between the two response channels) and the outcomes of interest—and not at the level of the individual cross-correlations, the non-stationarity of the original time series data is unproblematic.

Finally, there is the consideration of whether to base subjective ratings of experience on discrete emotions or broad affective dimensions. Although we often try to elicit “pure” (i.e., discrete) emotions in laboratory settings, emotion elicitors tend to produce complex blends of emotion or sequences in which one emotion segues into another (Levenson, 2003). Therefore, asking subjects to rate how strongly they feel a specific emotion from moment to moment risks underestimating the intensity of their overall subjective experience because it fails to capture blends of specific emotions on the same end of the valence continuum. Although this could in theory be addressed by having subjects provide continuous ratings for multiple discrete emotions simultaneously, such an approach would surely place an undue cognitive burden, causing a “competition of cues” (Pennebaker, 1982) and impinging on subjects’ natural affective and interoceptive trajectories. Asking subjects to rate the extent to which they feel a specific emotion from moment-to-moment might also bias them in the direction of the queried emotion category, producing ratings that more strongly reflect subjects’ exteroceptive judgments of the experimental stimuli than what subjects are feeling inside.

The alternative is to have subjects provide continuous ratings of their experience on broad affective dimensions, such as valence or arousal. Valence and arousal dimensions are thought to represent the basic aspects of semantic knowledge about emotion, an interpretation consistent with Osgood’s semantic differential work (Osgood, Suci, & Tannenbaum, 1957). Although such dimensions do not capture exhaustive information about emotion, they offer a useful tool for representing the core features of declarative knowledge about emotions (Kring, Barrett, & Gard, 2003). Of course, this raises the question of whether continuous ratings of subjective experience based on the valence dimension or the arousal dimension will track more closely with physiology within individuals; that is, which of these two rating dimensions furnishes a better proxy measure of interoceptive awareness? On the one hand, it could be argued that asking participants to tune into the sensation of their heartbeat and track the degree of physiological arousal in this channel from moment to moment (i.e., to provide arousal-based
ratings) amounts to a face-valid index of interoceptive awareness and is therefore the most straightforward way to assess this capacity. On the other hand, research shows that by the time visceral sensations become available to conscious awareness, they are already richly embedded with subjective valuations (i.e., positive/negative judgments), as evidenced by the activation of brain regions associated with hedonic valence (Craig, 2002; Craig, Chen, Bandy, & Reimann, 2000). From this latter perspective, valence-based ratings would appear to contribute to a more ecologically valid measure of interoceptive awareness. Because these competing perspectives have never been compared empirically, it seemed prudent to test both and directly compare them.

**Figure 1. Measuring Emotional Response Coherence as a Proxy for Interoceptive Awareness**

![Image of a diagram showing the trajectory of an emotional episode and the rationale for considering emotional response coherence as a proxy measure of interoceptive awareness. Included are the instruments (rating dial) and indices (heart period, somatic activity) used to measure different aspects of the emotional response package that are correlated in Coherence Scores.]

*Note.* This figure depicts the trajectory of an emotional episode and the rationale for considering emotional response coherence as a proxy measure of interoceptive awareness. Included are the instruments (rating dial) and indices (heart period, somatic activity) used to measure different aspects of the emotional response package that are correlated in Coherence Scores.

**Interoception and Distal Outcomes: Emotional Awareness, Psychological Adjustment, and Socio-emotional Sensitivity**

To determine whether coherence between subjective experience and physiology constitutes a psychometrically sound way to operationalize interoceptive awareness, its reliability and validity (convergent, predictive, discriminant) need to be established. Regarding predictive validity in particular, we would expect coherence to relate in similar ways to psychosocial outcomes that have been empirically linked to interoceptive awareness in the literature. I introduce some of these outcomes below.

Alexithymia is a personality construct characterized by impaired emotional awareness and deficits in the ability to identify and describe one’s feelings that can have profound effects on mental health and social functioning (Aleman, 2005; Bagby, Parker, & Taylor, 1994; Nemiah,
Interoception has been associated with alexithymia in numerous investigations (Bernhardt et al., 2014; Herbert, Herbert, & Pollatos, 2011; Hogeveen, Bird, Chau, Krueger, & Grafman, 2016; Kanbara & Fukunaga, 2016; Lemche et al., 2013; Wiebking & Northoff, 2015). For example, a large study including both genders (N=155) found that interoceptive awareness, assessed by a heartbeat counting task, was inversely related to all facets of the Toronto Alexithymia Scale (Herbert et al., 2011). Alexithymia appears to involve altered morphology and activation of the anterior insula and anterior cingulate cortex—two brain regions centrally involved in interoceptive processing (Berthoz et al., 2002; Borsci et al., 2009; Kano et al., 2003; Lane et al., 1998; Lane, Sechrest, Riedel, Shapiro, & Kaszniak, 2000). Alexithymia has also been associated with several psychopathologies posited to involve interoceptive impairments, including major depressive disorder (Bankier, Aigner, & Bach, 2001; Honkalampi, Hintikka, Transkanen, Lehtonen, & Viinamäki, 2000; Leweke, Leichsenring, Kruse, & Hermes, 2012; Saarijärvi, Salminen, & Toikka, 2001), autism (Bird et al., 2010), eating disorders (Brewer, Cook, Cardi, Treasure, & Bird, 2015; Monteboracci et al., 2006; Zonnevyle-Bender et al., 2005), and somatization and somatoform disorders (Burba et al., 2006; Karvonon et al., 2005).

Symptoms of anxiety, depression, and loneliness, which have a high prevalence in the general population, represent useful metrics for assessing psychological adjustment broadly, and all three have been examined in conjunction with interoception. Most prominently but also most problematically, interoceptive awareness has been implicated in the etiology and maintenance of anxiety disorders (Clark et al., 1997). Implicit in cognitive models of anxiety (e.g., Barlow, 1988; Beck, Emery, & Greenberg, 1985; Clark, 1986) is the assumption that individuals who are vulnerable to anxiety have a heightened propensity to not only perceive subtle changes in internal bodily sensations, but also to furnish them with dysfunctional cognitive appraisals characterized by threat-related interpretive biases. Consistent with this notion, positive associations between interoception and anxiety disorders—especially panic disorder—have been found in a large number of studies using both self-report questionnaires and heartbeat perception tasks (Dunn, Stefanovitch, et al., 2010; Ehlers & Breuer, 1992; Pollatos, Traut-Mattausch, & Schandry, 2009; Stevens et al., 2011; Zoellner & Craske, 1999). These findings should be interpreted with caution. As noted earlier, many self-report measures of body awareness were exclusively designed to assess anxiety and thus fail to disentangle simple awareness of body sensations from dysfunctional cognitive appraisals associated with these sensations (Mehling et al., 2009). As I underscored earlier, interoceptive awareness refers to the accurate perception of bodily states and is distinct from somatosensory amplification, exaggerated and noxious perceptions of somatic states that are a common feature of many anxiety disorders (Kanbara & Fukunaga, 2016). Indeed, in one review of the literature on interoception in anxiety, the authors concluded that anxiety is characterized not by interoceptive awareness per se but by an “altered interoceptive state resulting from amplified, self-referential interoceptive predictive beliefs” (Paulus & Stein, 2010). Lending credence to this idea, a study directly examining the relationship between interoceptive awareness and somatosensory amplification found higher levels of somatosensory amplification in poor heartbeat detectors than in good heartbeat detectors (Mailloux & Brener, 2002). Only two studies have examined the relationship between cardiac interoceptive awareness and anxiety symptoms in non-clinical samples, and of these, one found no relationship between heartbeat perception and trait anxiety (Steptoe & Vögele, 1992) and the other found a negative relationship between them (DePascalis, Alberti, & Pandolfo, 1984).
Depression is to a large extent characterized by somatic symptoms including changes in appetite and weight, disturbed sleep, and sexual dysfunction (American Psychiatric Association, 2013; Beck, 1967). In addition, a variety of non-specific somatic complaints such as fatigue, weakness, dizziness, headaches, and pain (Jain, 2009; Kapffhamer, 2006; Simon, VonKorff, Piccinelli, Fullerton, & J., 1999) are primary indicators of depression across many cultures (Kim, 2010; Kleinman, 2004; Simon, VonKorff, Piccinelli, Fullerton, & Ormel, 1999; Yusim et al., 2010). Depression also frequently co-occurs with medically vexing somatic syndromes, such as irritable bowel syndrome, non-ulcer dyspepsia, fibromyalgia, chronic fatigue, and chronic pain (den Boeft et al., 2016; Gatchel, Peng, Peters, Fuchs, & Turk, 2007; Henningsen, Zimmermann, & Sattel, 2003; Lépine & Briley, 2004). It is not surprising, then, that recent evidence increasingly points to the centrality of the body in depression, and more specifically, to interoceptive abnormalities (for a review, see Harshaw, 2015). Multiple behavioral and psychophysiological studies have reported poorer heartbeat perception in depressed individuals than in healthy controls (Dunn, Dagleish, Ogilvie, & Lawrence, 2007; Dunn, Stefanovitch, et al., 2010; Furman, Waugh, Battacharjee, Thompson, & Gotlib, 2013; Pollatos et al., 2009; Terhaar, Viola, Bär, & Debener, 2012). One of these studies found this to be concomitant with reduced heartbeat-evoked potentials in depressed individuals (Terhaar et al., 2012), suggesting that the neural activity underlying interoception may be altered in depression.

The insula is one of the primary cortical structures underlying interoceptive processing and awareness. One group of researchers (Paulus & Stein, 2006, 2010) has theorized that the way in which interoceptive afferents are integrated with representations of self in the insula might contribute to the pathogenesis of both anxiety and depression. According to their model, biased beliefs (i.e., propositional statements about the individual’s state that are processed in the medial prefrontal cortex and temporal-parietal junction with which the insula is connected) influence the evaluation of anticipatory interoceptive signals. Specifically, a tendency to exaggerate valence (esp. negative valence) amplifies the aversive aspects of predictive body signals, generating the anticipation of aversive bodily states. This increase in “background noise” reduces the signal-to-noise ratio when processing interoceptive information, making it harder for a person to differentiate between internal body signals that are associated with potentially aversive or pleasant consequences and those that are part of the ongoing and fluctuating internal milieu. Over time and through conditioning, afferent interoceptive signals (e.g., heartbeat, respiration) become imbued with catastrophic appraisals (e.g., “there’s something wrong with my heart”). This relative over-activity of cognitive control brain regions results in an increased production of thoughts and beliefs; practically, it is experienced as “worrying” aimed at improving prediction accuracy (Paulus & Stein, 2010). Lending empirical support to this model, one fMRI study comparing unmedicated depressed adults and healthy controls during an interoceptive attention task (Avery et al., 2014) found that the depressed group had less bilateral activity in dorsal mid-insula cortex than the non-depressed group when attending to interoceptive signals (i.e., heartbeat, stomach), and that greater task-related activity in the insula was associated with less severe depressive and somatic symptom symptoms in the depressed group. Another study (Farb, Segal, & Anderson, 2013) found that mindfulness meditation training contributes to interoceptive awareness-related functional plasticity by (a) promoting greater functional connectivity between the posterior and anterior insula, leading to better propagation of the interoceptive signal, and (b) reducing recruitment of dorsomedial prefrontal cortex and its connectivity with the insula, leading to diminished conceptual cortical activity and enhanced interoceptive activity (i.e., less “noise”).
Perceived loneliness has been linked to elevated levels of sympathetic nervous system activity, hypothalamic-pituitary-adrenal axis activity, and pro-inflammatory cytokines, and to downregulation of antiviral gene expression (Cole et al., 2015; Hawkley & Cacioppo, 2010)—all of which increase the risk of chronic disease and mortality (Holt-Lunstad, Smith, & Layton, 2010; Luo, Hawkley, Waite, & Cacioppo, 2012). Short-term body awareness training (i.e., mindfulness meditation) has been found to reduce self-reported loneliness; to ameliorate stress-related elevations in blood pressure, cortisol, and anxiety (Hughes et al., 2013; Tang et al., 2007); to downregulate loneliness-related pro-inflammatory gene expression and circulating protein biomarkers of inflammation (Cresswell et al., 2012; Tang et al., 2007); and to increase immunoreactivity (Tang et al., 2007). At the neural level, body awareness training appears to enhance cerebral blood flow to the anterior insula and anterior cingulate cortex—two key regions involved in interoceptive processing (Tang, Lu, Feng, Tang, & Posner, 2015). Interoceptive awareness, assessed by heartbeat perception tasks, has been found to attenuate negative affective responses to social exclusion, a likely precursor to perceived loneliness (Pollatos, Matthias, & Keller, 2015; Werner, Kerschreiter, Kindermann, & Duschek, 2013).

Broadly, interoception is thought to constitute the sense of self and to shape the way we experience the world (Berlucchi & Aglioti, 2010; Craig, 2002, 2009, 2010; Critchley et al., 2004; Park & Tallon-Baudry, 2014; Varela, Thompson, & Rosch, 1991). It is critical for emotional awareness (Dunn et al., 2007; Herbert et al., 2011; Silani et al., 2008) and emotion regulation (Füstös, Gramman, Herbert, & Pollatos, 2013; Kever et al., 2015; Koch & Pollatos, 2014); guides decision-making (Dunn, Galton, et al., 2010; Furman et al., 2013; Lamm & Singer, 2010; Sanfey et al., 2003; Singer et al., 2009; Sütterlin, Schulz, Stumpf, Pauli, & Vögele, 2013); and shapes self-control behaviors that have an impact on health and disease (Herbert, Blechert, Hautzinger, Matthias, & Herbert, 2013; Herbert, Herbert, et al., 2012; Herbert & Pollatos, 2014; Herbert et al., 2007). Given this swath of processes to which interoception is vital, we would expect interoceptive awareness to be positively associated with overall psychological wellbeing. This notion is supported indirectly by recent work (Lewis, Kanai, Rees, & Bates, 2014) that finds a positive association between gray matter volume in the right insular cortex, a region intimately involved in interoception, and self-reported psychological wellbeing assessed by the Ryff Scales (Ryff & Keyes, 1995).

A growing literature suggests that interoceptive awareness is also involved in social processing. This makes good sense given the central role of interoception in emotion, and the intimate association between emotional and social processing. Studies examining the associations between interoceptive awareness and social processing have focused primarily on empathy. In one study, individuals with greater interoceptive awareness, assessed by a heartbeat counting task, were found to be more sensitive to others’ emotional facial expressions (Terasawa, Moriguchi, Tochizawa, & Umeda, 2014). Another study (Fukushima, Terasawa, & Umeda, 2011) used heartbeat-evoked potentials, a neural measure of interoception that assesses cortical processing of cardiac activity, while subjects completed a task involving a mix of empathic judgment trials (i.e., rating the valence of emotional facial expressions in photos) and control trials (i.e., rating the symmetry of eyes on the faces in photos). Whereas heartbeat evoked potentials differentiated between the two trial types, raw cardiac measures (i.e., heart rate and EKG waveforms) did not. Moreover, the amplitude of heartbeat evoked potentials during empathic judgment trials was positively correlated with the Empathic Concern subscale of the Interpersonal Reactivity Index (Davis, 1983). Finally, the anterior insula repeatedly emerges in
the literature as a region involved in both interoceptive and empathic processing (e.g., Lamm & Singer, 2010; Singer et al., 2009).

The Present Study

The present study sought to examine whether the coherence between ratings of subjective experience and objective heart period during emotional episodes provides a novel way to assess interoceptive awareness. I have argued that the “Coherence Task” comprises specific features that ipso facto make it a proxy measure of interoceptive awareness, and further, a measure that is more ecologically meaningful and sensitive to individual differences than our ones. I briefly summarize these features below.

First, the Coherence Task assesses within-individual response coherence between continuous subjective experience and autonomic physiology. Second, heart period is the physiological signal examined in conjunction with subjective experience because (a) the heart is a powerful source of visceral sensation, (b) heart period is affected by both positive and negative emotion, and (c) existing studies of interoceptive awareness have focused primarily on heart rate and heart period. Third, time-lagged cross-correlations are used to assess the within-individual relationship between subjective experience and autonomic physiology because this approach accounts for the temporal dynamics and non-stationarity of time-series data. Fourth, valence-based and arousal-based momentary ratings of subjective experience are both examined and compared because it is not clear which will produce a better measure of interoceptive awareness. Finally, response coherence is examined while subjects view film stimuli that (a) reliably elicit autonomic activation, visceral sensations, and emotional responses in the lab (Gross & Levenson, 1995; Hubert & de Jong-Meyer, 1990; Levenson, 2003; McHugo, Smith, & Lanzetta, 1982); (b) induce variable levels of emotional arousal; and (c) sample both the positive and negative ends of the valence spectrum.

Having developed this Coherence Task, I evaluate its psychometrics as a proxy measure of interoceptive awareness using a series of tests. First, I assess its temporal stability by having subjects complete the task at two time points spaced one week apart. Second, I assess its convergent validity by examining the association between Coherence Task performance and a self-report measure of body awareness. Third, I assess its predictive validity by examining the associations between Coherence Task performance and emotional awareness (Alexithymia), psychological adjustment (i.e., anxiety, depression, loneliness, wellbeing), and socio-emotional sensitivity (i.e., empathy)—constructs that have been linked to interoceptive awareness in past research. Fourth, I assess its discriminant validity by directly comparing the coherence between subjective experience and heart period with an alternate version of the task in which somatic activity is used in lieu of heart period as the physiological index. Somatic activity has a timescale, continuity, and lack of error/artifact comparable to heart period, but represents a proprioceptive aspect of the emotional response package rather than an interoceptive one. If the Coherence Task is a proxy measure of interoceptive awareness as I am proposing, it should have specificity for the relationship between subjective experience and a visceral, or interoceptive signal (i.e., heart period) over a motoric or proprioceptive one (i.e., somatic activity). Finally, to explore the argument that interoceptive awareness is distinct from somatic amplification, which involves anxious cognitions that actually detract from the benefits of interoceptive awareness, I tested whether trait anxiety moderated the relationship between interoceptive awareness (assessed via the Coherence Task) and psychological wellbeing.
Method

Participants. Fifty-six adult men and women aged 18 to 50 were recruited through the Research Participation Program in the Psychology Department at the University of California, Berkeley. A power analysis using G-power (Faul, Erdfelder, Buchner, & Lang, 2009) indicated this sample size provided adequate power (.80) for detecting a medium effect size ($\beta = 0.32$) when computing two-tailed linear bivariate regressions (i.e., hypothesis testing at the between-person level examining the relation between Coherence Scores [cross-correlations] and outcomes of interest) using an alpha level of .05.

Apparatus and Measures

I. Measure of Response Coherence (Proxy for Interoceptive Awareness)

Subjective emotional experience. In line with previous work examining within-subject response coherence (Mauss et al., 2005; Sze et al., 2010), participants used an affect rating dial (Ruef & Levenson, 2007) to provide continuous ratings of their subjective emotional experience. They did so while watching montages of film clips designed to elicit a range of affective states with varying valence and intensity. During each of the two sessions spaced one week apart, participants completed two trials. In the first, they provided valence-based continuous ratings of subjective experience while their autonomic physiology was measured; in the second, they provided arousal-based continuous ratings of subjective experience while their physiology was measured. For the valence trials, the affect rating dial had a pointer that traversed 180-degrees distributed over nine divisions ranging from “Very Negative” (−4) to “Neutral” (0) to “Very Positive” (+4). For the arousal trials, the same affect rating dial was relabeled to range from “Least Aroused” (1) to “Most Aroused” (9). A computer sampled the rating dial position every 5 milliseconds and averaged these readings into 1-second measurement periods. The intensity of participants’ valence-based ratings of subjective experience was computed as the magnitude of the displacement of the rating dial position from the midpoint (0, neutral) in either the negative or the positive direction because sympathetic activation of the cardiovascular system can be affected by both intense negative and intense positive emotion (Bradley & Lang, 1997). The intensity of participants’ arousal-based ratings of subjective experience was computed as the displacement of the rating dial position from the lowest position (i.e., 1, Least Aroused).

Film stimuli. Emotion-eliciting film stimuli were designed to induce dynamic changes in affective state on both valence and arousal dimensions. Because all participants completed four trials in all (to compare valence and arousal ratings [i.e., Trial 1 versus Trial 2, within the same session] and to assess test-retest reliability [Session 1 and Session 2, spaced 1-week apart]), and the risk of habituation associated with such repetition, four different film montages were created to enable participants to view a novel montage during each trial. All four film montages were designed to be the same length (7.5 minutes) and to elicit the same sequence of emotions—sadness, nurturant love, disgust, calm, and strong negative arousal, in that order. Presentation of these film montages was counterbalanced within and across the two sessions to control for the potential effects of film viewing order. Details of these four montages are as follows. Montage 1: A lonely elderly man hangs himself after being released from prison (sadness), baby animals...
frolic as gentle lullaby-type music plays in the background (nurture love), a man forces himself to choke down a “milkshake” of squirming maggots and flies (disgust), fish swim under the ocean in aesthetically pleasing configurations with instrumental music in the background (calm), one man assaults another by kicking his head against a curb (strong negative arousal). Montage 2: A woman’s daughter is ripped from her arms by a Nazi soldier to be taken to her death (sadness), a second version of the baby animals frolicking clip using novel footage (nurture love), a man defecates into a filthy toilet then sifts through his feces looking for a package of drugs (disgust), a serene landscape with instrumental accompaniment (calm), a child in the slums is forced by gang leaders to shoot his young friend (strong negative arousal). Montage 3: A boy cries after his father dies in a boxing match (sadness), a third version of the baby animals frolicking (nurture love), a woman eats “spaghetti” made of worms and blood balls while repeatedly gagging (disgust), a different underwater fish scene (calm), a woman having her frenulum pierced screams in horror during the process (strong negative arousal). Montage 4: A woman reacts to news from a emergency room doctor that her two children have died (sadness), a fourth version of the baby animals frolicking (nurture love), a dog poops then the owner picks up and eats the feces (disgust), a different serene landscape scene (calm), a woman screams hysterically in the moments before she is hung to death (strong negative arousal).

Heart period. Continuous heart period, the interval (in msec) between successive R-waves on the electrocardiogram (EKG), was obtained using a Biopac polygraph, a computer with analog-to-digital capability and an online data acquisition and analysis computer program written by Robert W. Levenson. Two EKG electrodes were placed on the participant’s torso in a bipolar configuration. The same computer that acquired the rating dial data also acquired the heart period data, and both measures were averaged into 1-second periods. Although heart period was the physiological index of focal interest in the computation of coherence, a range of physiological responses were monitored, including impedance cardiography, finger pulse transmission time, ear pulse transmission time, respiration period, respiration depth, blood pressure, skin conductance, and somatic activity (used in discriminant validity analyses). Specifically, somatic activity was measured continuously via a pressure sensor under the participant’s chair using the same computer that acquired the rating dial data. The coherence between subjective experience and somatic activity was directly compared to that between subjective experience and heart period to assess the latter’s specificity for interoceptive awareness.

Coherence Score Calculation. Scores of the coherence between subjective experience (rating dial position; valence, arousal) and objective physiology (heart period or somatic activity) during film viewing were computed for each individual using a lagged cross-correlational analysis following procedures similar to those used in past research (Mauss et al., 2005; Sze et al., 2010). For the valence trials, the maximum cross-correlation coefficient between rating dial position (displacement from the neutral midpoint in either the negative or positive direction) and heart period (or somatic activity) within a 6-second lag window (i.e., −6 to +6) for each participant served as the within-individual measure of coherence. In light of previous evidence of age- and culture-based differences in the valuation of high-arousal versus low-arousal states (Scheibe, English, Tsai, & Carstensen, 2013; Tsai, Knutson, & Fung, 2006), I assessed the valence trials using both the signed cross-correlation coefficient and the absolute value of the cross correlation coefficient of the coherence between ratings of subjective experience and
physiology (heart period, somatic activity). The signed cross-correlation coefficient captures the maximum correlation between these two channels of responding for each individual, positive or negative, at whichever lag this occurred from \(-6\) to \(+6\); in other words, the signed cross-correlation coefficient specifies directionality. For example, a subject who experiences the physical sensations associated with increased sympathetic activation as unpleasant would move the dial toward the more negative end of the dial when his/her heart period decreased producing a negative cross-correlation coefficient, whereas a subject who experiences increased sympathetic activation as pleasant would move the dial toward the more positive end when his/her heart period decreased producing a positive cross-correlation coefficient. By contrast, the absolute value of the cross-correlation coefficient captures the maximum coherence between subjective experience and heart period, without heed to the direction of this correlation, at whichever lag this occurred from \(-6\) to \(+6\). By comparing both of these coherence indices for the valence trials, I was able to account for possible age- and culture-based differences in the valuation of physiological arousal.

For the arousal trials, the maximum cross correlation coefficient between the rating dial and heart period (or somatic activity) data within a 6-second lag window for each participant served as the coherence measure. The 6-second lag window was chosen because it conforms to theoretical notions about the duration and temporal characteristics of the different emotional subsystems whose coherence was assessed. Specifically, whereas heart period and somatic activity are part of the initial emotional response and thus generated more rapidly, subjective emotional experience is constructed afterward (Levenson, 1999, 2011) and thus would theoretically lag behind.

II. Outcome Measures for Validity Testing

**Self-Reported Interoception.** The Body Awareness Questionnaire (Shields et al., 1989) assesses awareness of a range of bodily processes (e.g., “I notice distinct body reactions when I am fatigued”). I administered a modified 13-item version of this measure, developed and used in past research in the Berkeley Psychophysiology Lab (Sze et al., 2010), to compensate for the fact that although items on the original measure assess awareness of visceral sensations in various contexts, they neglect to consider the context of emotion. In the interest of precision and parsimony, five of the original 18 items (i.e., those focusing on simple awareness of sensations and devoid of content reflecting somatosensory amplification, psychological distress, or pain) were included verbatim, while eight items were re-worded to capture body awareness associated with emotions and specific physiological signals (e.g., “I feel a distinct set of physical sensations occurring throughout my body when I feel sad as opposed to angry”; “I can often feel my heart beating”; “When I am feeling an emotion, I am not often aware of physical changes or sensations occurring in my body” [reverse-scored]). Respondents were asked to rate the accuracy of each statement using a Likert scale from 1 (Not at all true of me) to 7 (Very true of me).

**Emotional Awareness.** The Toronto Alexithymia Scale (Bagby, Parker, et al., 1994) is a widely used instrument comprising three subscales reflecting factor-driven facets of alexithymia (Nemiah et al., 1976; Taylor et al., 1999): difficulties identifying feelings (7 items, e.g., “I am often confused about what emotion I’m feeling”); difficulty describing feelings (5 items, e.g., “It is difficult for me to find the right words for my feelings); and externally oriented thinking (8 items, e.g., “I prefer talking to people about their daily activities rather than their feelings”).
used this instrument to assess emotional awareness. Respondents read 20 statements and rate the extent to which they agree with each using a Likert scale from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*). The Toronto Alexithymia Scale has been found to have good internal consistency, test-retest reliability, and convergent, discriminant, and concurrent validity (Bagby, Taylor, & Parker, 1994).

**Anxiety.** The trait subscale of the Spielberger State-Trait Anxiety Index (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) assesses the general tendency toward anxiety. Respondents read 20 statements describing feelings of anxiety and worry (e.g., “I worry too much over something that doesn’t really matter”; “I am calm, cool, and collected” [reverse-scored]) and use a Likert scale to rate the degree to which they generally feel this way ranging from 1 (*Not at all*) to 4 (*Very Much So*). This measure has good internal consistency and test-retest reliability, as well as considerable evidence attesting to its construct and concurrent validity (Spielberger, 1989; Spielberger et al., 1983).

**Depression.** The Center for Epidemiological Studies-Depression Scale (Radloff, 1977), was designed to assess depressive symptomatology in the general population rather than in clinical samples. Respondents read 20 statements describing various symptoms of depression (e.g., sad mood, difficulty concentrating, loss of appetite, insomnia, etc.) and use a Likert scale to rate how often they felt that way during the past week from 1 (*Rarely or None of the Time*) to 4 (*Most or All of the Time*). The CESD has good test-retest reliability ($r=0.87$; Radloff, 1977).

**Loneliness.** The UCLA Loneliness Scale (Russell, 1996) is a widely used measure that assesses subjective feelings of loneliness or social isolation. Respondents read 20 statements and rate the frequency with which they experience the feelings described using a Likert scale ranging from 1 (*Never*) to 4 (*Always*). Half of the statements are worded in a negative or “lonely” direction (e.g., “How often do you feel that there is no one you can turn to?”) and half are worded in a positive or non-lonely direction (e.g., “How often do you feel that you are ‘in tune’ with the people around you?”). This scale has good internal consistency (alphas ranging from .89 to .94), 1-year test-retest reliability ($r = .73$), and convergent validity (Russell, 1996).

**Psychological Wellbeing.** The Ryff Scales of Psychological Wellbeing-Short Form (Ryff & Keyes, 1995) is a 42-item inventory—20 positively worded items and 22 negatively worded (reverse-scored) items—assessing psychological wellbeing in six theoretically based domains: Autonomy (e.g., “I am not afraid to voice my opinions, even when they are in opposition to the opinions of most people”); Environmental Mastery (e.g., “In general, I feel I am in charge of the situation in which I live”); Personal Growth (e.g., “I think it is important to have new experiences that challenge how you think about yourself and the world”); Positive Relations with Others (e.g., “Maintaining close relationships has been difficult and frustrating for me,” reverse-scored); Purpose in Life (e.g., “I enjoy making plans for the future and working to make them a reality”); and Self-Acceptance (e.g., “In many ways, I feel disappointed about my achievements in life,” reverse scored). Respondents rate each of these statements using a Likert scale from 1 (*Strongly Disagree*) to 6 (*Strongly Agree*). All six subscales have good test-retest reliability (.81-.88), and the scale was found to negatively predict multiple dimensions of psychological distress one year later in a large sample of adults (N=1,179; Abbot et al., 2006).
**Empathy.** The Interpersonal Reactivity Index (Davis, 1983) comprises four 7-item subscales designed to assess various facets of dispositional empathy defined as “the reactions of one individual to the observed experiences of another.” The personal distress subscale is distinct from the other three in that it assesses “self-oriented” feelings of personal anxiety and unease in charged interpersonal scenarios (e.g., “Being in a tense emotional situation scares me”); the empathic concern subscale assesses “other-oriented” feelings of sympathy and compassion for unfortunate others (e.g., “I often have tender, concerned feelings for people less fortunate than me”); the fantasy subscale assesses the tendency to imaginatively transpose oneself into the feelings and actions of fictitious characters (e.g., “I really get involved with the feelings of the characters in a novel”); and the perspective-taking subscale assesses the tendency to spontaneously adopt the psychological point of view of others in everyday life (e.g., “When I'm upset at someone, I usually try to "put myself in his shoes" for a while”). Respondents read 28 statements and rate how accurately each describes them using a Likert scale from 0 (Does not describe me well) to 4 (Describes me very well). The four subscales demonstrate acceptable internal consistency, construct validity, and discriminant and convergent validity (De Corte, Buysse, Verhofstadt, & Roeyers, 2007).

**Procedure**

Each participant completed two lab-based experimental sessions conducted at the same time of day exactly one week apart. Participants had completed the self-report questionnaires online one to three days prior to their first laboratory session. On arriving at the Berkeley Psychophysiology Laboratory for their first session, participants were informed: “We are interested in examining links between body sensations and emotional experience.” Participants then entered the experimental room (a well-lit, 10 × 20-foot space), where they had physiological recording sensors attached to their bodies and were seated in a chair. Once all physiological signals were obtained without artifacts, participants were oriented to the affect rating dial. During the valence trial, which all participants completed first in both sessions, the rating dial was labeled from Very Negative to Neutral to Very Positive. Participants were instructed to “Move the dial to indicate how negative or positive you feel from moment to moment while watching the film.” During the arousal trial, the same rating dial was relabeled from Least Aroused to Most Aroused. Participants were instructed to “Tune into the sensation of your heart and, based on the sensation of your heartbeat, move the dial to indicate how physically aroused, or activated, you feel from moment to moment during the film.”

Emotional stimuli (film montages) were presented on a 27-inch color monitor positioned at a distance of 5.5 feet from the participant. Based on a computer-generated pseudo-randomization scheme, participants were assigned to one of four film sequences (A, B, C, or D), which prescribed the order in which they would view the four film montages across the two experimental sessions. Each film montage was preceded by a 1-minute baseline, during which participants were instructed to watch an ‘X’. Again, while viewing the first film montage, participants provided continuous valence-based subjective ratings and while viewing the second montage, participants provided continuous arousal-based subjective ratings. Participants repeated these tasks during the second session while viewing two novel film montages, enabling me to assess test-retest reliability. Counterbalancing the order in which participants viewed the four film montages allowed me to rule out any potential influences of film order on the outcome of interest.
Hypotheses

Hypothesis 1: Test-Retest Reliability. Given my thesis that coherence between subjective experience (valence, arousal) and heart period constitutes a proxy measure of interoceptive awareness, and because interoceptive awareness is argued to be trait-like and relatively stable over time, I hypothesized that Coherence Scores (i.e., cross-correlation coefficients derived from the coherence between subjective experience and heart period) would demonstrate stability as evidenced by significant 1-week test-retest reliability.

Hypothesis 2: Convergent Validity. Given my thesis that coherence between subjective experience (valence, arousal) and heart period constitutes a proxy measure of interoceptive awareness, I hypothesized that Coherence Scores would be positively associated with self-reported interoceptive awareness, indexed by the Body Awareness Questionnaire.

Hypothesis 3. Predictive Validity. Given empirical evidence for associations of interoceptive awareness with emotional awareness (alexithymia), psychological adjustment (anxiety, depression, loneliness, wellbeing), and socio-emotional sensitivity (empathy), I hypothesized that Coherence Scores would show a similar pattern of associations with these social-emotional outcomes. Specifically, I hypothesized that greater coherence between subjective experience (valence, arousal) and heart period would be associated with better emotional awareness (i.e., negatively correlated with alexithymia); psychological adjustment (i.e., negatively correlated with anxiety, depression, and loneliness; positively correlated with wellbeing), and socio-emotional sensitivity (i.e., positively correlated with empathy).

Hypothesis 4. Discriminant Validity. Given my thesis that coherence between subjective experience (valence, arousal) and heart period constitutes a proxy measure of interoceptive awareness, I expected the aforementioned hypotheses to have specificity for Coherence Scores derived from the coherence between subjective experience and an interoceptive physiological signal (i.e., heart period) over those derived from the coherence between subjective experience and a proprioceptive physiological signal (i.e., somatic activity). Specifically, I hypothesized that the coherence between subjective experience (valence, arousal) and heart period would demonstrate stronger test-retest reliability, convergent validity (i.e., association with self-reported interoceptive awareness), and predictive validity (i.e., association with outcomes linked to interoceptive awareness in past research) than would the coherence between subjective experience and somatic activity.

Hypothesis 5. The Moderating Effect of Trait Anxiety. I argued that the tendency to overlay threat-related cognitive interpretations onto perceived body signals is distinct from the simple awareness of interoceptive information, and that these cognitive appraisals are likely to detract from any positive psychological benefits conferred by having good interoceptive awareness. Accordingly, I hypothesized that Trait Anxiety (assessed by the Spielberger Trait Anxiety Index) would moderate the relationship between interoceptive awareness (indexed by the Coherence Task) and psychological wellbeing (indexed by the Ryff Scales); specifically, that Trait Anxiety would negatively moderate this relationship, such that individuals with higher levels of Trait Anxiety would show weaker associations between Coherence and Psychological Wellbeing.
Results

Preliminary analyses indicated that all dependent variables were normally distributed according to established cut-offs of an absolute value of 2 for skewness and kurtosis (Field, 2009; Gravetter & Wallnau, 2014; Trochim & Donnelly, 2006). Participant characteristics are presented in Table 1. All participants had complete data for demographic characteristics, Coherence Score and outcome measures used in hypothesis testing.

Table 1. Participant Characteristics

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>56</td>
<td>22.7 (5.8)</td>
</tr>
<tr>
<td>Gender (% Female)</td>
<td>56</td>
<td>61</td>
</tr>
<tr>
<td>BMI</td>
<td>56</td>
<td>23.6 (4.4)</td>
</tr>
<tr>
<td>Smoker (%)</td>
<td>56</td>
<td>14.3</td>
</tr>
<tr>
<td>Self-Reported Interoceptive Awareness (BAQ)</td>
<td>56</td>
<td>64.3 (11.7)</td>
</tr>
<tr>
<td>Alexithymia (TAS)</td>
<td>56</td>
<td>44.7 (11.6)</td>
</tr>
<tr>
<td>Anxiety (STAI-T)</td>
<td>56</td>
<td>43.3 (12.8)</td>
</tr>
<tr>
<td>Depression (CESD)</td>
<td>56</td>
<td>37.1 (9.9)</td>
</tr>
<tr>
<td>Loneliness (UCLA)</td>
<td>56</td>
<td>44.1 (10.8)</td>
</tr>
<tr>
<td>Psychological Wellbeing (Ryff)</td>
<td>56</td>
<td>181.3 (28.2)</td>
</tr>
<tr>
<td>Personal Distress (IRI)</td>
<td>56</td>
<td>12.8 (4.8)</td>
</tr>
<tr>
<td>Perspective Taking (IRI)</td>
<td>56</td>
<td>19.1 (4.8)</td>
</tr>
<tr>
<td>Fantasy (IRI)</td>
<td>56</td>
<td>19.2 (5.5)</td>
</tr>
<tr>
<td>Empathic Concern (IRI)</td>
<td>56</td>
<td>19.7 (4.5)</td>
</tr>
</tbody>
</table>

Note. BAQ = Body Awareness Questionnaire; STAI-T = Spielberger State Trait Anxiety Index-Trait subscale; CESD = Center for Epidemiological Studies Depression scale; UCLA = UCLA Loneliness Scale; Ryff = Ryff Scales of Psychological Wellbeing; IRI = Interpersonal Reactivity Index.

Test-Retest Reliability (Hypothesis 1). Because interoceptive awareness is posited to be a trait-like and relatively stable over time (Garfinkel & Critchley, 2013), my first analysis tested the hypothesis that coherence (proposed to be a proxy for interoceptive awareness) would demonstrate temporal stability. I also sought to examine which particular combination of dimensions—i.e., subjective ratings (valence vs. arousal) and objective physiology (heart period vs. somatic activity) would produce the Coherence Score(s) with the strongest 1-week test-retest reliability. I decided a priori to exclude any combination(s) of dimensions that did not demonstrate significant test-retest reliability from subsequent analyses. To test Hypothesis 1, I conducted a series of Pearson bivariate correlations between Coherence Scores (i.e., cross-correlation coefficients produced by the various combinations of subjective ratings and physiology) obtained during Session 1 and Coherence Scores obtained during Session 2, performed one week later. To address previous evidence of age- and culture-based differences in the valuation of high-arousal versus low-arousal states (Scheibe et al., 2013; Tsai et al., 2006), I examined both the signed cross-correlation coefficient (i.e., maximum positive cross correlation coefficient of the coherence between subjective responses and physiology) and the absolute
value of the cross correlation coefficient (i.e., cross correlation coefficient of the coherence between subjective experience and objective physiology, irrespective of the sign of the correlation) for Coherence Scores derived from valence-based subjective ratings. Results of these analyses are presented in Table 2. Because none of the Coherence Scores derived from the absolute value of the cross-correlation coefficients demonstrated significant reliability, I excluded these variables from all subsequent analyses.

Table 2. One-Week Test-Retest Reliability of Coherence Scores

<table>
<thead>
<tr>
<th>Parameters Used to Calculate Coherence Score</th>
<th>N</th>
<th>Correlation (Session 1 and 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence-Based, Heart Period, Absolute Value</td>
<td>56</td>
<td>( r = .26, p = .958 )</td>
</tr>
<tr>
<td>Valence-Based, Heart Period, Signed Maximum</td>
<td>56</td>
<td>( r = .47, p = .000^* )</td>
</tr>
<tr>
<td>Arousal-Based, Heart Period, Signed Maximum</td>
<td>56</td>
<td>( r = .33, p = .012^* )</td>
</tr>
<tr>
<td>Valence-Based, Somatic Activity, Absolute Value</td>
<td>56</td>
<td>( r = .08, p = .553 )</td>
</tr>
<tr>
<td>Valence-Based, Somatic Activity, Signed Maximum</td>
<td>56</td>
<td>( r = .20, p = .148 )</td>
</tr>
<tr>
<td>Arousal-Based, Somatic Activity, Absolute Value</td>
<td>56</td>
<td>( r = .51, p = .000^* )</td>
</tr>
</tbody>
</table>

Note. Parameters include subjective rating dimension, physiological channel, and method for calculating cross-correlation coefficient. * = Coherence Scores with significant test-retest reliabilities, established a priori as necessary for inclusion in subsequent hypothesis testing.

Preliminary Analyses for Validity Testing

Testing for Film-Specific Effects. As mentioned earlier, to circumvent the risk of habituation associated with repeated exposures to the same stimulus, I created four different film montages enabling participants to view a novel film montage during each of the four trials completed across the two experimental sessions. To rule out the possibility of film-specific effects (i.e., systematic differences in the degree to which the film montages elicited positive/negative emotions and/or arousal) on the Coherence Scores, I ran a one-way ANOVA with group (i.e., film montage viewing sequence to which participants were pseudo-randomly assigned: A, B, C, or D) as the predictor and the eight Coherence Scores\(^2\) as dependent variables. Results of these analyses are presented in Table 3 and indicate that the four film montages were statistically indistinguishable from each other in terms of their effect on the Coherence Scores.

\(^2\) Eight Coherence Scores were compared to determine which constituted the most psychometrically sound proxy for interoceptive awareness: the signed maximum cross correlation coefficient of the coherence between valence-based subjective ratings and heart period (Sessions 1 and 2); the signed maximum cross correlation coefficient of the coherence between arousal-based subjective ratings and heart period (Sessions 1 and 2); the signed maximum cross correlation coefficient of the coherence between valence-based subjective ratings and somatic activity (Sessions 1 and 2); and the signed maximum cross correlation coefficient of the coherence between arousal-based subjective ratings and somatic activity (Sessions 1 and 2).
### Table 3. Tests of Potential Film-Specific Effects on Coherence Scores

<table>
<thead>
<tr>
<th>Parameters Used to Calculate Coherence Score</th>
<th>Comparison (ANOVA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session 1, Valence Ratings, Heart Period</td>
<td>$F(3,52) = .63, p = .602$</td>
</tr>
<tr>
<td>Session 2, Valence Ratings, Heart Period</td>
<td>$F(3,52) = 1.14, p = .343$</td>
</tr>
<tr>
<td>Session 1, Arousal Ratings, Heart Period</td>
<td>$F(3,52) = 1.09, p = .362$</td>
</tr>
<tr>
<td>Session 2, Arousal Ratings, Heart Period</td>
<td>$F(3,52) = 2.31, p = .087$</td>
</tr>
<tr>
<td>Session 1, Valence Ratings, Somatic Activity</td>
<td>$F(3,52) = 1.04, p = .384$</td>
</tr>
<tr>
<td>Session 2, Valence Ratings, Somatic Activity</td>
<td>$F(3,52) = 1.40, p = .255$</td>
</tr>
<tr>
<td>Session 1, Arousal Ratings, Somatic Activity</td>
<td>$F(3,52) = 2.00, p = .126$</td>
</tr>
<tr>
<td>Session 2, Arousal Ratings, Somatic Activity</td>
<td>$F(3,52) = 2.04, p = .092$</td>
</tr>
</tbody>
</table>

*Note.* Parameters include session, subjective rating dimension, and physiological channel.

**Interrelations Among Candidate Coherence Scores.** Before directly comparing the eight candidate Coherence Scores on psychometrics to determine which serves as the best proxy for interoceptive awareness, I conducted Pearson bivariate correlations among them. The resulting zero-order correlation matrix is presented in Table 4. I studied the overall pattern of interrelations among these Coherence Scores looking for expected associations and to rule out any aberrant patterns. Because somatic activity is the strongest driver of heart period, I expected to see significant associations between the Coherence Scores derived using heart period and those derived using somatic Activity for the same Trial Type (i.e., valence or arousal) and Session (1 or 2). Results indicated that this assumption was met by all but one pair of Coherence Scores (i.e., that derived from coherence between arousal-based subjective ratings and heart period/somatic activity during Session 2), and that the associations between heart period-based and somatic activity-based Coherence Scores for the same trial and session were strongest for the Session 1 arousal trial and Session 2 valence trial. Although I did not exclude any Coherence Scores on this basis, the analysis highlighted some potentially problematic ones and provided a rough framework for understanding the subsequent analyses.
Table 4. Zero-Order Correlations Among the Eight Coherence Variables

<table>
<thead>
<tr>
<th></th>
<th>S1 Val, IBI</th>
<th>S2 Val, IBI</th>
<th>S1 Arous, IBI</th>
<th>S2 Arous, IBI</th>
<th>S1 Val, Act</th>
<th>S2 Val, Act</th>
<th>S1 Arous, Act</th>
<th>S2 Arous, Act</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1 Val, IBI</td>
<td>Corr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>.470**</td>
<td>.136</td>
<td>.085</td>
<td>.310*</td>
<td>.129</td>
<td>.287*</td>
<td>.359**</td>
</tr>
<tr>
<td>N</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>S2 Val, IBI</td>
<td>Corr</td>
<td>.470**</td>
<td>1</td>
<td>.350**</td>
<td>.233</td>
<td>.401**</td>
<td>.432**</td>
<td>.274*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.000</td>
<td>.008</td>
<td>.085</td>
<td>.002</td>
<td>.001</td>
<td>.041</td>
<td>.018</td>
</tr>
<tr>
<td>N</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>S1 Arous, IBI</td>
<td>Corr</td>
<td>.136</td>
<td>.350**</td>
<td>1</td>
<td>.332*</td>
<td>.101</td>
<td>.409**</td>
<td>.581**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.318</td>
<td>.008</td>
<td>.012</td>
<td>.458</td>
<td>.002</td>
<td>.000</td>
<td>.021</td>
</tr>
<tr>
<td>N</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>S2 Arous, IBI</td>
<td>Corr</td>
<td>.085</td>
<td>.233</td>
<td>.332*</td>
<td>1</td>
<td>.184</td>
<td>.001</td>
<td>.172</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.532</td>
<td>.085</td>
<td>.012</td>
<td>.176</td>
<td>.996</td>
<td>.204</td>
<td>.100</td>
</tr>
<tr>
<td>N</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>S1 Val, Act</td>
<td>Corr</td>
<td>.310*</td>
<td>.401**</td>
<td>.101</td>
<td>.184</td>
<td>1</td>
<td>.196</td>
<td>.277*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.020</td>
<td>.002</td>
<td>.458</td>
<td>.176</td>
<td>.148</td>
<td>.039</td>
<td>.357</td>
</tr>
<tr>
<td>N</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>S2 Val, Act</td>
<td>Corr</td>
<td>.129</td>
<td>.432**</td>
<td>.409**</td>
<td>.001</td>
<td>.196</td>
<td>1</td>
<td>.442**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.343</td>
<td>.001</td>
<td>.002</td>
<td>.996</td>
<td>.148</td>
<td>.001</td>
<td>.001</td>
</tr>
<tr>
<td>N</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>S1 Arous, Act</td>
<td>Corr</td>
<td>.287*</td>
<td>.274*</td>
<td>.581**</td>
<td>.172</td>
<td>.277*</td>
<td>.442**</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.032</td>
<td>.041</td>
<td>.000</td>
<td>.204</td>
<td>.039</td>
<td>.001</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>S2 Arous, Act</td>
<td>Corr</td>
<td>.359**</td>
<td>.316*</td>
<td>.309*</td>
<td>.222</td>
<td>.126</td>
<td>.416**</td>
<td>.514**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.007</td>
<td>.018</td>
<td>.021</td>
<td>.100</td>
<td>.357</td>
<td>.001</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
</tr>
</tbody>
</table>

Note. Val = valence-based subjective ratings, Arous = arousal-based subjective ratings, IBI = interbeat interval, Act = somatic activity, Corr = Pearson correlation, Sig = significance, * = significant at the 0.05 level (2-tailed), ** = significant at the 0.01 level (2-tailed).

**Convergent Validity (Hypothesis 2).** To examine the association between coherence and self-reported interoceptive awareness, and to determine which of the eight candidate Coherence Scores demonstrated the strongest association with subjective interoceptive awareness, I conducted a multiple regression analysis with the Body Awareness Questionnaire as the dependent variable and the eight Coherence Scores as predictors. I considered whether to include Age, Gender, and BMI as covariates in this analysis, given their robust associations with interoceptive awareness in the literature (e.g., Jones, 1994; Rouse et al., 1988; Wiens & Palmer, 2001; Yates et al., 1985) by running bivariate correlations of Age, Gender, and BMI with the Body Awareness Questionnaire. Results (Table 5) revealed that none was significantly associated with self-reported interoceptive awareness, so I did not include them as covariates in the subsequent analysis.
Table 5. Associations of Age, Gender, BMI with Self-Reported Interoceptive Awareness

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Correlation with Body Awareness Questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>56</td>
<td>$r = .17$, $p = .207$</td>
</tr>
<tr>
<td>Gender</td>
<td>56</td>
<td>$r = .15$, $p = .260$</td>
</tr>
<tr>
<td>BMI</td>
<td>56</td>
<td>$r = .20$, $p = .134$</td>
</tr>
</tbody>
</table>

In the multiple regression with the Body Awareness Questionnaire as the outcome and the eight candidate Coherence Scores as predictors, only two Coherence Scores came out as significant predictors: the Coherence Score derived from valence-based subjective ratings and heart period during Session 2 was a significant positive predictor, $b = .33$, $t(55) = 2.08$, $p = .043$, and that derived from arousal-based subjective ratings and heart period during Session 2 was a significant negative predictor, $b = -.48$, $t(55) = -3.57$, $p = .001$. Because the latter association was negative (i.e., in the direction opposite to that hypothesized, indicating poor convergent validity), I did not examine it further. All other $ps$ were $\geq .308$. Next, I ran a follow-up regression with the Body Awareness Questionnaire as the dependent variable and the Coherence Score derived from valence-based subjective ratings and heart period during Session 2 (i.e., the only significant predictor in the preceding analysis) to see if it explained significant variance in self-reported interoceptive awareness. Results showed that the Coherence Score derived from valence-based subjective experience and heart period explained 7.8% of the variance in the Body Awareness Questionnaire, which was significant, $F(1,54) = 4.46$, $p = .039$.

**Predictive Validity (Hypothesis 3).** To assess whether coherence relates in predictable ways to emotional awareness, psychological adjustment, and socio-emotional sensitivity—outcomes linked to interoceptive awareness in the literature, I followed a series of steps. First, in the interest of parsimony and to minimize Type I Error, I conducted a factor analysis with Trait Anxiety (Spielberger State-Trait Anxiety Index, Trait subscale), Depression (Center for Epidemiological Studies-Depression Scale), Psychological Wellbeing (Ryff Scales), Loneliness (UCLA Loneliness Scale), and Empathy (Interpersonal Reactivity Index (IRI), four subscales). Results are shown in Table 6 and Figure 2. The analysis yielded a two-factor solution composed of an intrapersonal “Distress” factor with five loadings—Anxiety, Depression, Psychological Wellbeing (negative), Loneliness, and IRI Distress—and an interpersonal “Empathy” factor with three loadings—IRI Empathy, Fantasy, and Perspective-Taking subscales.
Table 6. Factor Analysis of Psychosocial Functioning Measures, Item Loadings

<table>
<thead>
<tr>
<th></th>
<th>Component 1</th>
<th>Component 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety (STAI)</td>
<td>.937</td>
<td></td>
</tr>
<tr>
<td>Depression (CESD)</td>
<td>.889</td>
<td></td>
</tr>
<tr>
<td>Psychological Wellbeing (Ryff’s)</td>
<td>- .859</td>
<td>.327</td>
</tr>
<tr>
<td>Loneliness (UCLA-L)</td>
<td>.813</td>
<td>-.342</td>
</tr>
<tr>
<td>IRI Distress</td>
<td>.660</td>
<td></td>
</tr>
<tr>
<td>IRI Empathic Concern</td>
<td>-.109</td>
<td>.849</td>
</tr>
<tr>
<td>IRI Fantasy</td>
<td>.122</td>
<td>.755</td>
</tr>
<tr>
<td>IRI Perspective-Taking</td>
<td>-.198</td>
<td>.714</td>
</tr>
</tbody>
</table>


**Figure 2. Factor Analysis of Psychosocial Functioning Measures, Scree Plot**

Next, I computed two factor analytically derived composites, “Distress” and “Empathy” by z-scoring and averaging their respective component items, then performed a reliability analysis on each one. The Distress composite (5 items) had a Cronbach’s alpha of .90 and the Empathy composite (3 items) had a Cronbach’s alpha of .68.

For the main analyses of predictive validity to determine which of the eight Coherence Scores are the strongest predictors of psychosocial outcomes associated with interoceptive awareness, I conducted three separate multiple regressions with the Toronto Alexithymia Scale.
(Emotional Awareness), the factor analytically derived Distress composite, and the factor analytically derived Empathy composites, in turn, as dependent variables. As before, I first considered whether to include Age, Gender, and BMI as covariates by running bivariate correlations between these three variables and the three outcomes of interest (Table 7). Because Age, Gender, and BMI were not significantly associated with Alexithymia, Intrapersonal Distress, or Empathy, I did not include them as covariates in the subsequent regressions.

Table 7. Associations of Age, Gender, BMI with Emotional Awareness, Distress, & Empathy

<table>
<thead>
<tr>
<th></th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Toronto Alexithymia Scale</td>
</tr>
<tr>
<td>Age</td>
<td>N = 56</td>
</tr>
<tr>
<td>Gender</td>
<td>N = 56</td>
</tr>
<tr>
<td>BMI</td>
<td>N = 56</td>
</tr>
</tbody>
</table>

In the first regression testing the association between Coherence and Emotional Awareness, none of the eight candidate Coherence Scores was significantly associated with the Toronto Alexithymia Scale, all $bs \leq |.18|$, all $ps \geq .300$.

In the second regression testing the association between Coherence and Distress, the Coherence Score derived from Valence-Based Subjective Ratings and Heart Period during Session 2 emerged as a marginally negative predictor, $b = -.35$, $t(55) = -1.94$, $p = .059$. None of the remaining Coherence Scores was significantly associated with Distress, all $ps \geq .096$. I then ran a follow-up regression with Distress as the dependent variable and this Coherence Score as the predictor to see if it accounted for significant variance in Distress. Results showed that coherence explained 6.7% of the variance in Distress, which was marginally significant, $F(1,54) = 3.86, p = .055$. I reran the regression predicting Distress, this time including Age, Gender, and BMI as Step 1 covariates and the eight Coherence Scores as Step 2 predictors. The Coherence Score derived from valence-based subjective ratings and heart period during Session 2 emerged as the only significant predictor, but this time it was a significant negative predictor, $b = -.38$, $t(55) = -2.01$, $p = .049$. I then conducted a follow-up regression with Distress as the outcome, Age, Gender, and BMI as Step 1 covariates, and only this Coherence Score as a predictor in Step 2. The full model explained 11.9% of the variance in Distress, which was not significant, $F(4,51) = 1.72, p = 1.60$. Although the covariates accounted for 3.5% of the variance in Distress, which was not significant, $F(3,52) = .623, p = .603$, the Coherence Score explained an additional 8.4% of the variance in Distress, which was significant, $\Delta F(1,51) = 4.87, p = .032$.

In the third regression predicting Empathy, none of the eight Coherence indices emerged as a significant predictor, all $ps \geq .138$. When I reran the analysis including Age, Gender, and BMI entered as covariates in Step 1 and the eight Coherence Scores in Step 2, the Coherence Score derived from valence-based subjective ratings and heart period during Session 2 emerged
as a marginally positive predictor of Empathy, \( b = .32, t(55) = -1.82, p = .076 \), but the remaining Coherence Scores were non-significant, \( ps \geq .126 \).

I then conducted a follow-up regression with Empathy as the dependent variable and the Coherence Score derived from valence-based subjective ratings and heart period during Session 2 as the only predictor to see if it explained significant variance in Empathy. I ran this regression with and without the covariates; in both cases, neither the full model nor the Coherence Score itself explained significant variance in Empathy, all \( ps \geq .151 \).

**Discriminant Validity (Hypothesis 4).** As mentioned earlier, tests of discriminant validity (i.e., whether Coherence Scores derived from subjective ratings and heart period, a visceral signal, would be a better proxy measure of interoceptive awareness than a Coherence Scores derived from subjective ratings and somatic activity, a motoric signal) were embedded throughout the preceding analyses. Specifically, Coherence Scores derived from subjective ratings and heart period and those derived from subjective ratings and somatic activity were directly compared in the assessments of test-retest reliability, convergent validity, and predictive validity above. Although the Coherence Scores derived from subjective ratings and somatic activity showed good test-retest reliability (for arousal, but not for valence), none of the Coherence Scores derived using somatic activity as the physiological component demonstrated convergent validity (i.e., a significant association with self-reported interoceptive awareness) or predictive validity (i.e., significant associations with emotional awareness, distress, or empathy).

**The Moderating Effect of Trait Anxiety (Hypothesis 5).** To test my final hypothesis that trait anxiety would moderate the relationship between interoceptive awareness and psychological wellbeing, I conducted a stepwise multiple regression predicting Psychological Wellbeing using the Coherence Score that demonstrated the strongest psychometric properties in the previous analyses (i.e., the Coherence Score derived from valence-based subjective ratings and heart period during Session 2) as the predictor and Trait Anxiety as the moderator. To buttress support for this choice of Coherence Score, I also examined the bivariate correlations of all eight candidate Coherence Scores with Psychological Wellbeing (Table 8). Only the Coherence Score derived from valence-based subjective ratings and heart period during Session 2 was significantly positively associated with Psychological Wellbeing.

**Table 8. Associations of Candidate Coherence Variables with Psychological Wellbeing**

<table>
<thead>
<tr>
<th>Coherence Score</th>
<th>N</th>
<th>Pearson Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session 1, Valence-Based Ratings and IBI</td>
<td>56</td>
<td>( r = -.00, p = .976 )</td>
</tr>
<tr>
<td>Session 2, Valence-Based Ratings and IBI</td>
<td>56</td>
<td>( r = .27, p = .043^* )</td>
</tr>
<tr>
<td>Session 1, Arousal-Based Ratings and IBI</td>
<td>56</td>
<td>( r = .07, p = .592 )</td>
</tr>
<tr>
<td>Session 2, Arousal-Based Ratings and IBI</td>
<td>56</td>
<td>( r = -.20, p = .142 )</td>
</tr>
<tr>
<td>Session 1, Valence-Based Ratings and Activity</td>
<td>56</td>
<td>( r = .13, p = .328 )</td>
</tr>
<tr>
<td>Session 2, Valence-Based Ratings and Activity</td>
<td>56</td>
<td>( r = .25, p = .063 )</td>
</tr>
<tr>
<td>Session 1, Arousal-Based Ratings and Activity</td>
<td>56</td>
<td>( r = .15, p = .275 )</td>
</tr>
<tr>
<td>Session 1, Arousal-Based Ratings and Activity</td>
<td>56</td>
<td>( r = .23, p = .083 )</td>
</tr>
</tbody>
</table>

*Note. PWB = Psychological Wellbeing. * = statistical significance at the \( p < .05 \) level.*
Next, I considered whether to include Age, Gender, and BMI as covariates in the moderation analysis. Specifically, I ran bivariate correlations between these three variables and the outcome, Psychological Wellbeing (Table 9). Because Age, Gender, and BMI were not significantly associated with Psychological Wellbeing, I did not include them as covariates.

### Table 9. Associations of Age, Gender, BMI with Psychological Wellbeing

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Pearson Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>56</td>
<td>$r = .09, p = .494$</td>
</tr>
<tr>
<td>Gender</td>
<td>56</td>
<td>$r = .01, p = .958$</td>
</tr>
<tr>
<td>BMI</td>
<td>56</td>
<td>$r = .10, p = .453$</td>
</tr>
</tbody>
</table>

Finally, I ran the stepwise regression analysis predicting Psychological Wellbeing. In Step 1, I entered the centered predictor (Coherence Score) and moderator (Trait Anxiety) and in Step 2, I entered the interaction term (product of the two centered variables). Given the significant negative association between the predictor and moderator (correlation between Coherence Score and Trait Anxiety, $r = -.28, p = .037$), there was a risk of multicollinearity. I ruled this out by examining the Tolerance and Variance Inflation Factor (VIF)—specifically, all Tolerance values were $\geq .91$ and all VIF values were $\leq 1.10$, indicating that multicollinearity was not an issue. The main effects of Coherence and Trait Anxiety (Step 1) together explained 63.4% of the variance in Psychological Wellbeing, which was significant, $F(2,53) = 45.94, p = .000$. Although the main effect of Trait Anxiety was significant, $b = -.85, t(55) = -9.65, p = .000$, the main effect of Coherence was not, $b = -.01, t(55) = -.11, p = .917$. Adding the interaction term in Step 2 accounted for an additional 0.3% of the variance in Psychological Wellbeing, which was not significant, $\Delta F(1,52) = 0.37, p = .546$, suggesting that Trait Anxiety did not moderate the effect of Coherence on Psychological Wellbeing in this sample.

**Discussion**

In this dissertation, I presented a novel method that capitalizes on emotional response coherence to assess individual differences in interoceptive awareness. This method, the Coherence Task, assesses variability in the extent to which continuous physiological responses track with momentary ratings of subjective experience within individuals when they are experiencing strong emotions. This method yields a unique Coherence Score for every individual—namely, a cross-correlation coefficient representing the coherence between continuous heart period (a visceral signal) and continuous ratings of subjective experience—making it sensitive to the full spectrum of individual differences in emotional response coherence, and by implication, interoceptive awareness. This feature has eluded traditional heartbeat perception tasks, which categorize individuals as either ‘good heartbeat perceivers’ (rarely more than 40 percent of individuals) or ‘poor heartbeat perceivers,’ yielding dichotomous scores. A further advantage of this method is that it assesses coherence—and correspondingly, interoceptive awareness—in the face of emotionally evocative stimuli, which is how
interoceptive processes naturally unfold, rendering it a more ecologically valid way to assess interoceptive awareness than traditional heartbeat perception tasks.

Having grounded the Coherence Task in theoretically and empirically supported rationales, I evaluated its psychometric properties as a proxy measure of interoceptive awareness. Specifically, I examined its temporal stability (1-week test-retest reliability); convergent validity (association with self-reported interoceptive awareness); predictive validity (associations with outcomes linked to interoceptive awareness in the literature); and discriminant validity (specificity for coherence between subjective experience and a visceral or interoceptive physiological signal over a proprioceptive physiological one). I also examined which of two broad dimensions of emotion—valence or arousal—contributes to a more robust proxy measure of interoceptive awareness. Finally, to explore the argument that interoceptive awareness is distinct from somatic amplification, which involves anxious cognitions that actually detract from the benefits of interoceptive awareness, I tested whether trait anxiety moderated the relationship between interoceptive awareness (assessed via the Coherence Task) and psychological wellbeing.

In the test of temporal stability, scores derived from the coherence between both valence-based and arousal-based subjective ratings with heart period demonstrated significant 1-week test-retest reliability. However, the Coherence Score derived using valence-based subjective ratings yielded a slightly larger effect size and smaller p-value than did the one derived using arousal-based subjective ratings.

In the test of convergent validity, I compared eight Coherence Scores derived using all combinations of parameters (i.e., Session [1 and 2], subjective rating dimension [valence and arousal], and physiological channel [heart period and somatic activity]) to see which was most strongly associated with self-reported interoceptive awareness. Among the eight Coherence Scores, only the one derived from valence-based subjective ratings and heart period during Session 2 was significantly positively associated with the Body Awareness Questionnaire, and explained 7.8 percent of the variance in self-reported interoceptive awareness.

In the test of predictive validity, I compared the same eight Coherence Scores to see which was most strongly associated with outcomes linked to interoceptive awareness in the literature (i.e., emotional awareness, distress, and empathy). None of the Coherence Scores showed a significant association with emotional awareness. The score derived from the coherence between valence-based subjective ratings and heart period during Session 2 emerged as a significant negative predictor of Distress—a composite of anxiety, depression, perceived loneliness, interpersonal distress, and psychological wellbeing—and explained 8.4 percent of the variance. The Coherence Score derived using the same parameters was also a marginal positive predictor of empathy, a composite of the empathic concern, fantasy, and perspective-taking subscales of the Interpersonal Reactivity Index. Notably, the intra-item reliability for the Empathy composite was substantially lower than that for the Distress composite (Cronbach’s alphas of .68 and .90, respectively).

In terms of discriminant validity, the analyses above all included direct comparisons between scores based on the coherence between ratings of subjective emotional experience and heart period, a visceral signal, and scores based on the coherence between ratings of subjective emotional experience and somatic activity, a motoric signal. Accordingly, tests of specificity for interoceptive processing over proprioceptive processing (i.e., to determine whether the Coherence Task is a proxy measure of interoceptive awareness specifically rather than body awareness generally) were embedded. The fact that only one of the eight Coherence Scores (i.e.,
that derived from valence-based subjective ratings and heart period during Session 2) repeatedly emerged as the front runner in psychometric testing and was the only significant predictor of outcomes linked to interoceptive awareness suggests some specificity for interoceptive processing, or at the very least for heart period over somatic activity.

Finally, results did not support the hypothesis regarding the moderating effect of trait anxiety on the relationship between coherence and psychological wellbeing. Given the relatively small sample size and the fact that both the predictor and moderator were continuous variables, there is a good likelihood that this analysis was underpowered to detect interaction effects (McClelland & Judd, 1993). However, it is also the case that only the main effect of trait anxiety explained significant variance in psychological wellbeing; the main effect of Coherence Score did not.

Taken together, these results suggest that the Coherence Task—and more specifically, the cross-correlation coefficient of the coherence between valence-based subjective ratings and heart period obtained during Session 2—constituted a psychometrically sound proxy measure of interoceptive awareness, a finding that raises several questions.

First and foremost is whether the Coherence Task can be considered superior to traditional heartbeat perception tasks as an assessment of interoceptive awareness. Crucially, I did not compare the two methods directly in the present study, and thus cannot provide empirical support for the Coherence Task’s incremental validity. Several findings nonetheless provide indirect support for this notion. First, I established the temporal stability of this measure, which is important because interoceptive awareness is assumed to be biologically rooted and relatively stable within individuals over time (Garfinkel & Critchley, 2013); test-retest reliability of heartbeat discrimination tasks has not been established. Second, I used precisely the same Coherence Task when testing associations with self-reported interoceptive awareness (i.e., convergent validity) and with outcomes linked to interoceptive awareness in the literature (i.e., predictive validity). By contrast, studies of the associations between heartbeat perception tasks and these outcomes have involved a patchwork of heartbeat discrimination and heartbeat counting tasks, with specific methods varying greatly even within each of these task types. In other words, there is no single heartbeat perception task whose psychometric properties has been systematically established as I have done in the present investigation. Finally, because the coherence method developed and tested here captures a spectrum of individual differences in interoceptive awareness by generating a unique Coherence Score for every individual along a continuous distribution, it can be considered a more sensitive index than heartbeat perception tasks. Further, because the coherence method assesses interoceptive awareness under conditions involving emotionally salient and meaningful stimuli, it adds value beyond heartbeat perception tasks in terms of ecological validity.

Another finding that requires some interpretation is the superior performance of valence-based subjective ratings over arousal-based subjective ratings, even though it would seem that arousal-based ratings represent a more face-valid and straightforward way to assess interoceptive awareness. One possible clue as to why this was the case can be gleaned from the neuroscience of interoception. Research has shown that by the time visceral sensations are available to conscious awareness in the anterior insula, they have already become richly integrated with representations of hedonic state (Craig, 2002; Craig et al., 2000). Specifically, the posterior insula provides primary interoceptive information via topographically organized and modality-specific pathways to the anterior insula, where information about an individual’s interoceptive state and hedonic state are integrated through the anterior insula’s connections to corticolimbic
and striatal reward circuit components. These components include the hypothalamus, which maintains homeostasis in the internal milieu; the nucleus accumbens, which processes the incentive motivational aspects of rewarding stimuli (Reynolds & Zahn, 2005; Robinson & Berridge, 2008); the amygdala, which is involved in emotional arousal and is critical for processing stimulus salience, as well as emotional learning and memory (Augustine, 1985; Jasmin, Burkey, Granato, & Ohara, 2004; Jasmin, Rabkin, Granato, Boudah, & Ohara, 2003; Paton, Belova, Morrison, & Salzman, 2006; Reynolds & Zahn, 2005); the anterior cingulate cortex, which engenders motivational aspects of emotion and is involved in various tasks related to self-monitoring and evaluating action selection (Augustine, 1996; Critchley, Tang, Glaser, Butterworth, & Dolan, 2005; Goldstein et al., 2007; Reynolds & Zahn, 2005; Rushworth & Behrens, 2008); and the orbitofrontal cortex, which is implicated in the context-dependent evaluation of environmental stimuli (Bechara, Damasio, & Damasio, 2000; Kringelbach, 2005; O ’Doherty, Kringelbach, Hornak, Andrews, & Rolls, 2001; Ongür & Price, 2000; Rolls & Grabenhorst, 2008; Schoenbaum, Roesch, & Stalnaker, 2006; Schoenbaum, Setlow, Saddoris, & Gallagher, 2003). Accordingly, it makes sense that a Coherence Score derived from valence-based subjective ratings (i.e., one that accounts for hedonic judgments) would provide a functionally superior index of interoceptive awareness. As I argued in the introduction, research has tended to take an overly reductionist approach to measuring interoceptive awareness that fails to account for the emotional and motivational contexts in which interoceptive processing actually occurs. The insula is not only critical for sensing and mapping internal stimuli, it is also involved in evaluating and responding to the potential meaning and impact of these stimuli on the organism (Paulus & Stein, 2006). Indeed, the insula plays a crucial role in detecting emotionally salient stimuli (Morris et al., 1998; Phillips et al., 1998), and in generating and regulating affective responses (Phillips et al., 2003).

The explanation for the superior performance of the valence-based version of the Coherence Task might also be far simpler: in this study, subjects always completed the valence trials first (i.e., before the arousal trials). The decision not to counterbalance the valence and arousal trials was a deliberate one based on the rationale that participants tend to be fresher earlier in an experimental protocol. Because the present study builds on two previous ones that used valence-based subjective ratings (Levenson, Ekman, & Ricard, 2012; Sze et al., 2010), it seemed prudent to iterate on previous findings by testing valence-based subjective ratings under ideal conditions. At the same time, it leaves open the possibility that the superior psychometric performance of valence-based ratings over arousal-based ratings is simply attributable to the former being performed under more ideal conditions. I intend to address this issue directly in a replication by administering both trial types and counterbalancing them.

Another finding that requires some interpretation is that the significant associations between Coherence Scores (derived from valence-based subjective ratings and heart period) and outcomes relevant to interoceptive awareness were unique to Session 2. In other words, the same task completed during Session 1 did not yield similar associations with the outcomes of interest as might be expected, especially in light of the test-retest reliability. The simplest explanation for this discrepancy is that participants were better oriented to the Coherence Task during the Session 2. Making continuous subjective ratings using the dial while watching a film can be unnatural at first, requiring subjects to frequently look down at the dial and away from the film as they become acquainted with the dial positions until these become more automatic. This introduces subtle task-switching demands—i.e., the need to disengage from the film stimulus, orient to a new stimulus (rating dial), and then engage with the new stimulus (and vice versa),
again and again. This cognitive demand, which diminishes with practice, likely impeded subjects from being able to simultaneously immerse in the film, tune into their bodies, and report their experience from moment to moment, which likely mitigated the intensity of emotional responding—a key factor in emotional response coherence. As mentioned earlier, subjects completed the valence trials first in both sessions and did not complete any practice trials due to time constraints. During the Session 1 valence trial, the Coherence Task was completely novel to subjects, whereas by Session 2 subjects were relatively well acquainted with the rating dial having completed the Coherence task twice already during Session 1. If this explains the session-based discrepancy in performance, it is instructive in terms of refining the task protocol by including practice trials in future research.

**Limitations**

The present study had several limitations. The most prominent being that I did not directly compare the Coherence Task with conventional heartbeat perception tasks to establish incremental validity. As discussed in the introduction, heartbeat perception tasks are not very encouraging in terms of their psychometrics and ecological validity. In light of this fact, and considerations of time constraints and subject fatigue, it seemed more important to prioritize the inclusion of two types of subjective rating dimensions (i.e., valence and arousal). However, the stage is now set for a follow-up study that directly compares the Coherence Task with both heartbeat discrimination and heartbeat counting tasks to determine which is the strongest measure of interoceptive awareness, a project that is already underway.

Another limitation of this study is its sole focus on individual differences in awareness of a single visceral center: the heart. Interoceptive awareness involves multiple physiological and parameters, including other cardiac signals (e.g., cardiac output, obtained via impedance cardiography), blood pressure, respiratory rate and load, electrodermal activity, and gastric myoelectric activity. Moreover, individuals may vary in the particular physiological center to which they are most sensitive and which most strongly informs their subjective experience. If this is the case, then the Coherence Task presented here may not be assessing interoceptive awareness equally in all people, but rather, be biased toward individuals who are preferentially focused on and aware of their hearts. As I stated in the introduction, linguistic and psychophysiological studies suggest that the body-based metaphors we use to describe our emotions likely reflect our awareness of the underlying physiology and gut-centered metaphors (e.g., “gut feelings,” “stomach-churning,” etc.) are no less plentiful in our language than heart-centered ones. In terms of the Coherence Task, however, the temporal characteristics of gut physiology pose some constraint. Recall that it is imperative for two channels whose coherence is being examined to have the same temporal resolution: the affect rating dial and heart period both capture responses on a second-by-second basis. Gut responses, by contrast, unfold on a much slower timescale of three cycles on average on the electrocardiogram (EGG; Koch & Stern, 2004), making it difficult to assess their coherence with continuous ratings of subjective experience using the Coherence Task. An assessment of gut-based interoceptive awareness would thus require an altogether different protocol. To this end, I am currently gathering exploratory data that examines subjective awareness of EGG responses using the so-called “water load test,” a standardized, non-invasive test of gastric myoelectric activity that produces a reliable EGG response in healthy individuals. Because changes in gastric myoelectrical activity after a water load correspond to varying degrees with subjective perceptions of fullness (Herbert,
Muth, Pollatos, & Herbert, 2012; Koch & Stern, 2004), the water load test may provide a window into assessing individual differences in the subjective perceptions of objective gut activity (i.e., gut-based interoceptive awareness).

A final methodological limitation of this study centers on the discrepancy between the structural features of the rating dial I used for each of the two types of trials. Whereas the rating dial I used for the valence trials had a bipolar scale that increased in equal and opposite directions through a neutral midpoint (i.e., −4 being ‘Most Negative’, 0 being ‘Neutral’, and +4 being ‘Most Positive’), the dial I used for the arousal trials was unidirectional, going from 1 ‘Least Aroused’ to 9 ‘Most Aroused’. Given the nature of arousal, this makes conceptual sense: I considered subjects’ most relaxed state to represent their lowest level of arousal and any activation greater than this to represent an increase. Put another way, it is difficult to conceive what negative values or a “neutral” midpoint might mean in regard to the arousal dimension. It nonetheless remains possible that the discrepancy between these two rating dials confounded the comparison of rating dimension (i.e., valence vs. arousal) with the structural features of the scale (i.e., anchoring scheme and score computation).

**Implications and Future Directions**

The ability to understand and assess individual differences in interoceptive awareness is likely to have important implications across multiple fields of inquiry and practice, including clinical psychology, psychiatry, and neuroscience. In recent years, researchers and clinicians have noted a trend toward disembodiment, particularly within our Western techno-centric culture. This has become even more pronounced amid the surge of interest and reliance on monitoring the body using external devices, such as the Fitbit and Apple Watch, which shift attention and awareness away from internal signals and toward external ones. At the same time, there has been a growing trend in finding ways to return to the body and to cultivate present-moment awareness through practices as varied as yoga, meditation, dance, somatic psychotherapies, massage, cooking, craft-work, and convening with nature (Hassed, 2013; Leder, 1990; Mehling et al., 2009). Moreover, burgeoning public interest in stress reduction methods that draw on key aspects of interoceptive awareness (e.g., focused breathing, mindfulness) has stimulated vibrant inquiry and dialogue about mind-body relationships across areas as diverse as neuroscience, psychology, philosophy, and spirituality (Astin, Shapiro, Eisenberg, & Forsys, 2003; Barnes, Powell-Griner, McFann, & Nahin, 2004; NCCAM, 2004).

In regard to clinical psychology and psychopathology in particular, the field has historically placed a heavy emphasis on the interplays between environmental cues and cognitive-behavioral tendencies contributing to patterns of stimulus-response learning that condition maladaptive responses. For example, cognitive-behavioral theories highlight the way in which stimuli in the external environment trigger core beliefs and produce automatic thoughts that lead to dysfunctional behaviors in a cyclic fashion. But these models underplay the role of the internal environment in this cycle: cues in the environment are constantly engendering involuntary body responses (e.g., patterns of autonomic arousal) and generating powerful interoceptive feedback that directly influences feelings, judgments, and behavior in crucial ways (Nauta, 1971; Damasio, 1994). Consider, for example, how autonomic over-activation or under-activation creates sensations, hedonic judgments, and approach/avoidance motivations that shape an array of behaviors from drug and alcohol use, to physical activity, to risk-taking, to social withdrawal. It is thus not surprising that alterations in interoceptive awareness are increasingly
being implicated in a range of pathologies, including substance use disorders, chronic pain, mood and anxiety disorders, posttraumatic stress disorder, and eating disorders (Di Lernia, Serino, & Riva, 2016; Fischer et al., 2016; Harshaw, 2015; Khalsa et al., 2015; Lanius, Frewen, Tursich, Jetly, & McKinnon, 2015; Lattimore et al., 2017; Naqvi & Bechara, 2010; Paulus & Stein, 2006, 2010; Simmons, Strigo, Matthews, Paulus, & Stein, 2009; Verdejo-Garcia, Clark, & Dunn, 2012). Yet despite these very promising lines of inquiry, we still lack a standard protocol for assessing individual differences in interoceptive awareness that is both psychometrically sound and ecologically meaningful. If through further examination, replication, and refinement, the Coherence Task is found to meet these standards, it could make a valuable contribution to multiple disciplines. Not only would it provide an empirically grounded basis for understanding individual differences in interoceptive awareness and its functional correlates, it would also allow us to evaluate the efficacy of interventions that target interoceptive awareness in the service of improved health and wellbeing. Moreover, because a state of heightened emotional activation is presupposed by and “built into” the Coherence Task, it could be especially valuable to the study of pathologies characterized by heightened arousal and psychological inflexibility (e.g., posttraumatic disorder, substance use disorders, borderline personality disorder, and chronic pain). That is because individuals who are susceptible to rigid and compulsive behaviors (i.e., those who have difficulty integrating and responding to new information from the internal and external milieu from moment to moment) are particularly vulnerable during heightened states of emotional and physiological aroused.

To bolster our confidence that the coherence method developed and tested here does in fact serve as a proxy for interoceptive awareness, it will be necessary to build on the present study in several ways. First, to replicate the positive findings (i.e., those pertaining to reliability and convergent, predictive, and discriminant validity). In this replication, it will be essential to (a) counterbalance valence and arousal trials to determine whether the superior performance of valence-based trials found here merely stemmed from order effects or was due to a deeper association between hedonic judgment and interoceptive awareness; and (b) include practice trials so that participants are better oriented to using the rating dial and able to immerse in the film. Second, to establish incremental validity of the Coherence Task over conventional heartbeat perception tasks through a direct comparison with both heartbeat discrimination and heartbeat counting versions. Third, to compare other channels of visceral responding that occur on a second-by-second timescale (e.g., blood pressure, respiration, skin conductance); to include additional metrics within the same physiological channel (e.g., heart period versus cardiac output); and to examine physiological composites derived from combinations of these different channels. Finally, given the theoretical assumption that coherence between physiology and subjective experience is stronger in the context of emotional arousal, it will be informative to compare differences in Coherence Scores for the same person during high versus low emotional arousal states. Pursuing all of these avenues of inquiry would have been unwieldy in the scope of the present investigation, which aimed to establish “proof of concept” for the Coherence Task as a proxy measure of interoceptive awareness. However, my hope is that our further work examining the avenues outlined above will contribute to the development of a robust, sensitive, and ecologically meaningful way to assess individual differences in interoceptive awareness.
References


Lattimore, P., Mead, B., Irwin, L., Grice, L., Carson, R., & Malinowski, P. (2017). ‘I can’t accept that feeling’: relationships between interoceptive awareness, mindfulness and


