Behavior Research Methods
https://doi.org/10.3758/s13428-022-01817-4

Planning, conducting, and analyzing a psychophysiological experiment on challenge and threat: A comprehensive tutorial

Phillip R. Johnston[1,2] · **Alexandra E. Volkov**[3] · **William S. Ryan**[2] · **Spike W. S. Lee**[2,3]

Accepted: 22 February 2022

Abstract
The biopsychosocial model of challenge and threat (BPS-CT) is a powerful framework linking psychological processes to reliable patterns of cardiovascular responses during motivated performance situations. Specifically, the BPS-CT poses challenge and threat as two motivational states that can emerge in response to a demanding, self-relevant task, where greater challenge arises when perceived resources are higher than demands, and greater threat arises when perceived resources are lower than demands. By identifying unique patterns of physiological responses associated with challenge and threat, respectively, the BPS-CT affords insight into subjective appraisals of resources and demands, and their determinants, during motivated performance situations. Despite its broad utility, lack of familiarity with physiological concepts and difficulty with identifying clear guidelines in the literature are barriers to wider uptake of this approach by behavioral researchers. Our goal is to remove these barriers by providing a comprehensive, step-by-step tutorial on conducting an experiment using the challenge and threat model, offering concrete recommendations for those who are new to the method, and serving as a centralized collection of resources for those looking to deepen their understanding. The tutorial spans five parts, covering theoretical introduction, lab setup, data collection, data analysis, and appendices offering additional details about data analysis and equipment. With this, we aim to make challenge and threat research, and the insights it offers, more accessible to researchers throughout the behavioral sciences.

Keywords Challenge · Threat · Motivation · Psychophysiology · Cardiovascular processes

Introduction

The biopsychosocial model of challenge and threat (BPS-CT[1]; Blascovich & Mendes, 2000, 2010) is a framework for studying motivational states and their affective, cognitive, and situational antecedents. The model contrasts challenge and threat, two motivational states that emerge based on an individual's subjective appraisal of resources versus demands during a motivationally self-relevant and demanding situation or task. Despite their discrete labels, challenge and threat occupy opposite ends of a continuum, with greater challenge arising when perceived resources are higher than demands, and greater threat occurring when perceived resources are lower than demands (Blascovich, 2008; Blascovich & Tomaka, 1996; Seery, 2013).

The power of the BPS-CT comes from its basis in physiological processes. While both challenge and threat are characterized by increased physiological activation indicative of engagement in a demanding, self-relevant task, the specific patterning of these autonomic and neuroendocrine changes differs in identifiable and meaningful ways (Table 1). Tracking these patterns can provide insight into subjective appraisal of resources and demands, thereby providing a valuable complement to subjective measures, particularly where participants may be unwilling or unable to provide self-reports (Blascovich, 2008; Seery, 2013).

✉ Phillip R. Johnston
phil.johnston@mail.utoronto.ca

[1] Rotman Research Institute, Baycrest Centre, 3560 Bathurst Street, Toronto, ON M6A 2E1, Canada

[2] Department of Psychology, University of Toronto, 100 St. George Street, Toronto, ON M5S 3G3, Canada

[3] Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, ON M5S 3E6, Canada

[1] While 'BPS' has been used to denote the biopsychosocial model when discussing challenge and threat elsewhere in the literature (e.g. Blascovich et al., 2004; Jamieson et al., 2012; Seery et al., 2009), this tutorial uses BPS-CT throughout in order to distinguish the biopsychosocial model of challenge and threat specifically from other biopsychosocial models.

Published online: 23 May 2022

Table 1 Physiological profiles of challenge and threat

Cardiovascular Index	Reactivity associated with greater challenge (change from baseline to task)	Reactivity associated with greater threat (change from baseline to task)
Cardiac output (CO)	Increase	Decrease or no change
Total peripheral resistance (TPR)	Decrease	Increase or no change
Challenge and Threat Index (CTI)	Higher scores	Lower scores
Heart rate (HR)	Increase (prerequisite for task engagement)	Increase (prerequisite for task engagement)
Pre-ejection period (PEP)	Decrease (prerequisite for task engagement)	Decrease (prerequisite for task engagement)

Getting started with physiological indices and measurement techniques, however, can be daunting to those without a background in physiology or practical experience with physiological equipment. Indeed, existing resources often assume some level of physiological knowledge. Even those who are familiar with psychophysiology may be discouraged by a lack of centralized, concrete, and detailed information regarding methods (especially data processing and analysis) in the challenge and threat literature, because the relevant information that has been published is scattered over numerous journal articles and book chapters, posing a barrier to researchers hoping to incorporate these methods into their research program.

This paper aims to address these problems, with the goal of making challenge and threat research more accessible to researchers throughout the behavioral sciences, especially those with no prior knowledge or experience in psychophysiology. To that end, we provide a comprehensive, step-by-step tutorial on conducting a challenge and threat experiment from the design stage to data analysis, with detailed appendices and concrete examples. By the end of this tutorial, and selective review of the primary works referenced here, a researcher should have a solid foundation in challenge and threat research and cardiovascular psychophysiology more generally.

For those who are new to challenge and threat research, the tutorial can be read in the order presented. For those looking for a refresher on a particular area or direction to further resources, the topics are divided as follows.

Part I includes a brief review of the BPS-CT, its advantages and limitations, and applicability to various areas of behavioral research. This section also provides an overview of cardiovascular psychophysiology, its underlying logic, and the advantages and limitations of employing physiological measurements.

Part II provides a comprehensive guide to setting up your lab and choosing measurement hardware and software.

Part III discusses experiment design and elaborates on data collection.

Part IV covers data processing and analysis, and provides an overview of the various cardiovascular indices involved in challenge and threat research.

Appendix A lists all relevant acronyms. Appendix B provides supplementary information about data processing to complement software documentation. Appendix C offers a review of commonly needed items.

Part I: Cardiovascular psychophysiology and the biopsychosocial model of challenge and threat

What is psychophysiology and why is it useful?

Challenge and threat research belongs to the broader field of psychophysiology. Psychophysiologists study phenomena of psychological interest, such as affect, cognition, and motivation, as they are related to and revealed through physiological processes (Cacioppo et al., 2007). Unlike related fields such as psychobiology, psychophysiology emphasizes social, psychological, and behavioral processes, rather than lower-level physiological mechanisms.

In general, psychophysiological research (and thus challenge and threat research) operates under the central axiom that mental phenomena derive from the structure and activity of the brain and nervous system. In other words, mental processes are embodied phenomena, with measurable physiological correlates. Through careful experimentation and logic, psychophysiologists aim to learn about unobservable mental processes by measuring their attendant physiological variables (Cacioppo & Tassinary, 1990; Cacioppo et al., 2007).

Psychophysiologists focus on the physiological system that best suits their research question (Cacioppo et al., 2007). One such system is the cardiovascular system, which has been shown to carry rich information about motivational and stress-related processes, and is the system on which the BPS-CT centers (Blascovich & Mendes, 2000; Blascovich & Tomaka, 1996). Specifically, cardiovascular psychophysiology involves measuring or estimating parameters related to the heart (e.g., heart rate, stroke volume) and the vascular system (vasoconstriction/dilation) to index psychological processes. By examining changes in several cardiovascular indices (discussed below), challenge and threat can be

differentiated, allowing for the investigation of their social, cognitive, and affective antecedents.

Complementing self-report measures of subjective experience, well-defined, validated, and theoretically grounded physiological measures offer a number of practical benefits, owing to their *continuous*, *online*, and *covert* nature (Blascovich & Mendes, 2010). *Continuous* measures can capture bodily responses as they unfold over time, thereby providing insight into the dynamics of psychological experiences better than discrete self-report measurements. Relatedly, physiological responses are measured *online*, in real time, while subjective measures are often necessarily recorded before or after the event or process of interest has taken place. The BPS-CT and other psychophysiological models can therefore circumvent certain biases that arise when participants are asked to forecast, remember, or reflect on their experiences. Physiological measures also help overcome a related difficulty, in that they are sensitive to mental and physiological changes to which a participant may not have introspective access, and thus cannot report (Blascovich et al., 2011). Lastly, physiological measures are considered *covert* because participants typically do not actively monitor and adjust their responses.[2] This is particularly useful when investigating issues where impression management is a concern (e.g., intergroup interactions; Frings et al., 2012; Mendes et al., 2002; Mendes, Gray, et al., 2007).

Despite these advantages, it is important to note that physiological measures are not inherently better indices of psychological constructs than any other. Rather, their value depends on the theoretical foundation and empirical work used to establish their validity. Cacioppo and colleagues (1990; 2007) offer a comprehensive framework for understanding the strength of the relationship between psychological variables and physiological responses, which in turn determine the strength of inference that can be drawn about the former from the latter. In the ideal case, there would be an invariant (one-to-one) relationship between variables in the psychological and physiological domains, but in practice researchers must make use of less definitive, but still useful, relationships between index and construct. In the case of challenge and threat research specifically, the BPS-CT aims to tighten the relationship between physiological index and psychological construct by limiting the inferential context to motivated performance situations, and using a pattern of multiple physiological measures. We direct readers to Cacioppo et al. (2007) for more information about the logic of psychophysiological inference in general, and Blascovich et al. (2011) for details pertaining to the BPS-CT specifically.

[2] Although research does suggest that meditation and breathing techniques can impact physiological reactions via activation of the vagus nerve (e.g., Ditto et al., 2006), participants are typically unaware of or unlikely to employ such techniques unless instructed to do so as part of the experimental design.

Autonomic and neuroendocrine influences in cardiovascular psychophysiology

To better understand cardiovascular psychophysiology, and the threat and challenge constructs, it is important to have some fundamental knowledge of the major physiological systems involved, namely the autonomic nervous system and the neuroendocrine system.

The autonomic nervous system (ANS) is the body's primary regulatory system, controlling a wide range of physiological processes including heart rate, respiration, and digestion. Because the ANS orchestrates bodily changes in response to a variety of mental processes, it is a primary system of interest not only in cardiovascular psychophysiology, but psychophysiology in general.

The two subsidiary branches of the ANS, the sympathetic and parasympathetic nervous systems, have diverse functions and interactions throughout the body.[3] In instances of heightened motivation and arousal, sympathetic activity predominates, resulting in the "fight-or-flight" response, in which bodily responses oriented towards mobilizing resources occur, such as increased heart rate and raised blood glucose levels. This response is largely orchestrated by a sub-system of the sympathetic branch, known as the sympathetic-adrenal-medullary (SAM) axis, which controls the release of epinephrine from the adrenal medulla. With respect to the cardiovascular system, its main effects are to increase heart rate, blood pressure, and the contractile force of the heart. This is a fast-acting system, capable of reaching peak activation within seconds of the eliciting stimulus or mental event.

The SAM axis is not the only system involved in the body's response to high motivation and arousal. The hypothalamic–pituitary–adrenal (HPA; also known as pituitary–adrenal–cortical or PAC) axis, one of the body's primary neuroendocrine systems, modulates the body's response by releasing cortisol from the adrenal cortex. This cortisol release may inhibit epinephrine-mediated vasodilation, leading to higher blood pressure when the HPA and

[3] The sympathetic and parasympathetic nervous systems were originally defined anatomically, and a succinct description of their function is difficult given their various effects on end organ activity (Jänig, 2006). Similarly, it is worth noting that the commonly held notion of broadly antagonistic effects between the two systems misrepresents the fact that relatively few tissues are innervated by both systems, and modes of interactions vary in those that are (see Sect. 1.3 in Jänig, 2006). Antagonism does occur in the cardiovascular system (e.g., heart rate, contractility, and vasodilation/vasoconstriction), but this does not mean that sympathetic and parasympathetic control is reciprocal, where only one system dominates at a given time. Rather, autonomic cardiovascular control is better described by a bivariate model, allowing for sympathetic and parasympathetic activation to covary positively, negatively, or vary independently (Berntson et al., 1993, 1994a, 1994b; Cacioppo et al., 1994).

SAM axes are co-activated (Blascovich, 2008). In contrast to the SAM axis, these measurable effects of HPA activation and subsequent cortisol release unfold more slowly. It should be noted, however, that a long period of exposure is not a requirement for HPA axis activation. In other words, both HPA and SAM responses can be elicited by a brief period of high motivation and arousal.

In general, stress responses of this kind can be considered either adaptive or maladaptive, depending on their duration, intensity, and overall effect on the well-being of the individual, and activity of the HPA axis is considered key to this distinction (McEwen, 1998, 2004). While HPA activity can be adaptive in the short term (e.g., by promoting the mobilization of glucose in the bloodstream), heightened and/or prolonged HPA activity has been implicated in a variety of negative outcomes, including increased allostatic load, depression, and cardiovascular disease (Juster et al., 2010; McEwen, 1998, 2004).

The biopsychosocial model of challenge and threat (BPS-CT)

Challenge and threat theory emerged directly from Dienstbier's (1989) work on physiological toughness in rodents and Lazarus and Folkman's (1984) cognitive-appraisal theory. Dienstbier (1989) found that rodents that thrived in threatening environments exhibited activation of the SAM axis, whereas those that languished demonstrated both SAM and HPA activation. Although researching rodents, Dienstbier saw a connection between his findings and the way Lazarus and Folkman (1984) were theorizing about stress in humans – specifically that there are clear individual differences in patterns of responding to potentially stressful stimuli. In their cognitive-appraisal theory or transactional model of stress, Lazarus and Folkman and colleagues (Folkman et al, 1986; Lazarus & Folkman, 1984) emphasized (as the name suggests) that stress emerges from a complex transaction between the individual and environment. In other words, we can understand individual differences in responses to potentially stressful stimuli as following from individual differences in cognitive appraisals applied to the situation. The first, or *primary,* appraisal is whether this stimulus is self-relevant. In other words, is this stimulus harmful, challenging, or threatening to *me*? Next, a secondary appraisal process unfolds in which the individual evaluates the demands of the situation and whether they have the coping resources to meet these demands.

The BPS-CT integrates and builds on both of these research traditions while containing several key departures. Drawing on Lazarus & Folkman's work, the BPS-CT similarly considers the role of both the situation and the individual in determining patterns of physiological reactivity. The BPS-CT translates the neuroendocrine patterns identified by Dienstbier into the human cardiovascular context, identifying reliable patterns of reactivity (described below) indicative of the relative activation of the SAM and HPA axes and their accompanying psychological states, termed challenge and threat.

The BPS-CT differs from the transactional model of stress in several important ways. Whereas Lazarus & Folkman discuss appraisals as conscious cognitive meaning making processes, the BPS-CT considers appraisal processes to occur automatically without deliberation (Blascovich, 2008; Seery, 2011; Weisbuch-Remington, et al., 2005). At times, the term "evaluations" is used to emphasize this difference. The BPS-CT also explicitly limits the scope of its evaluation to "active coping" (Blascovich, 2008, 2013), or coping within contexts that are not only self-relevant, but require a cognitive and/or behavioral response on the part of the individual (Obrist, 1981). Thus, the cardiovascular indices of challenge and threat must be assessed within the context of what is termed a "motivated performance situation," one in which the condition of self-relevance is met and an active response is required. In other words, the task must induce cognitive and/or behavioural engagement toward a goal that is perceived as relevant to the participant. Such engagement is termed *task engagement*, can be indexed through cardiovascular measures, and is a physiological prerequisite for relative challenge or threat. Common motivated performance tasks designed to induce task engagement include preparing and giving a short speech (e.g., on why one would be a good friend or employee), completing word or visual puzzles, and mathematical computations. Given the evaluative nature of laboratory experiments, whether conducted in the presence of real or imagined others, most performance based tasks are likely to qualify as motivated performance situations with the potential to elicit task engagement. The BPS-CT, however, does not apply to "passive coping" (Obrist, 1981) situations, which do not require a response to the potentially stressful stimulus (e.g., watching a video).

The theories also differ in their use of the terms challenge and threat; while the transactional model considers challenge and threat to be part of the primary appraisal process contributing to the appraisal of self-relevance, the BPS-CT considers these to be the outcomes of the evaluations – first of self-relevance and the need for action (e.g., are the conditions of task engagement met?) and then of the ratio of resources relative to demands in the motivated performance situation. Demands may include danger, uncertainty, novelty, and required effort (Blascovich & Mendes, 2000). Resource evaluations involve relevant knowledge and abilities, dispositional characteristics (e.g., self-esteem, optimism), and external support. What is key is the resulting ratio, or relative balance, of these demands to resources, which may be influenced by subconscious stimuli (Weisbuch-Remington et al., 2005), personality characteristics

(Tomaka & Magoc, 2021), and social evaluation processes (Mendes, et al., 2001).

These evaluations, and the resulting ratio, are automatic, dynamic, and iterative rather than static (e.g., Quigley, et al., 2002); as the situation itself and/or one's evaluation of it changes, so will autonomic nervous system activation. Reflecting this dynamic and relative ratio of situational demands to resources, the challenge and threat model can best be thought of as a continuum lying between these two poles, rather than as dichotomous and discrete psychophysiological states (Seery, 2013).

To summarize, the BPS-CT posits that, when faced with a self-relevant situation that requires an active response (a motivated performance situation), individuals make automatic evaluations or appraisals as to the demands of the situation and the resources they have available to meet those demands. When resources are appraised as high relative to demands, a motivational state of challenge emerges. When situational demands are appraised as low relative to demands, a motivational state of threat emerges. According to the BPS-CT, these motivational states can be indexed by the specific patterns of reactivity in the autonomic and neuroendocrine systems already described. Specifically, the challenge end of the continuum is associated with increased activation of the SAM axis, whereas threat is marked by activation of both the SAM and HPA axes.

In order to infer the relative activation of these two systems, and thereby distinguish challenge and threat (assuming participants are adequately engaged in the task), the BPS-CT employs peripheral measures of cardiovascular reactivity[4] focusing on changes in cardiac efficiency and vascular resistance. Challenge is characterized by adaptive patterns of cardiovascular reactivity (similar to those observed when engaging in cardiovascular exercise) including increased cardiac output (CO) and decreased total peripheral resistance (TPR). This means that the heart is pumping more blood per minute and that the veins and arteries have expanded to accommodate this additional blood flow. Threat, however, is characterized by little change in CO and an increase in TPR (Table 1; Blascovich & Mendes, 2010). Little change in both measures in the presence of task engagement may also be consistent with threat (Wormwood et al., 2019). Because the vessels do not expand to accommodate faster and more forceful heart contractions, vascular resistance (TPR) can increase, but the overall amount of blood pumped by the heart (CO) does not increase.

The BPS-CT thus offers a means of distinguishing between different high arousal states using multiple measures (Blascovich, 2008; Blascovich & Kelsey, 1990; Blascovich & Tomaka, 1996; Cacioppo & Tassinary, 1990), and the utility of this approach for psychological and behavioral research has been demonstrated by its application across a variety of phenomena. Social psychological examples include intergroup relations (Mendes et al., 2002), stereotype threat (Hoyt & Blascovich, 2010; Vick et al., 2008), and interactions with stigmatized others (Blascovich et al., 2001). Intrapsychic examples include predicting future performance (Behnke & Kaczmarek, 2018; Blascovich et al., 2004; Seery et al., 2010), expectancy violation (Mendes, Blascovich, et al., 2007), reappraisal and framing (Jamieson et al., 2012; Seery et al., 2009), and decision making (Kassam et al., 2009).

In sum, the BPS-CT presents a flexible, validated, and relatively accessible approach to investigating motivational states across a wide variety of different motivated performance situations. Our goal is to make the practical aspects of this approach more readily achievable for everyone.

Part II: Lab setup, hardware, and software

Parts II and III guide the reader through the process of conducting a study using the BPS-CT, laid out in chronological order from lab setup to data collection. Appendix C lists all suggested items to acquire.

First considerations

Employing cardiovascular electrophysiology in your research program may involve difficulties not typically encountered when using other behavioral research methods. Certain factors should therefore be considered at the outset, in order to determine whether this methodology is appropriate for your lab or research question.

First, continuous, online physiological measurement requires specialized equipment and software, which necessitates a more significant financial investment than do typical self-report and behavioral measures. While prices vary between suppliers, a full setup can cost upwards of several thousand US dollars, and reach much higher if continuous blood pressure measurement is desired (see Blood Pressure (BP) below). In addition to this initial investment, there are also the recurring costs of electrodes (if using disposable electrodes), conductive gel, and other consumables to consider. While some of these costs can be reduced somewhat by sharing equipment amongst collaborators, funding is still a necessary consideration.

[4] Note that measuring cardiovascular activity is not the only way of indexing autonomic and neuroendocrine activities relevant to the BPS-CT. Some researchers may be interested in measuring cortisol and alpha-amylase, which can be used to infer HPA axis and ANS activity, respectively, albeit at much longer sampling intervals than cardiovascular methods (Ali & Nater, 2020; Blascovich et al., 2011; Cacioppo et al., 2007; Kirschbaum & Hellhammer, 1989).

In addition to monetary concerns, getting started with psychophysiological methods requires an investment of time and energy. Beyond that necessary to familiarize oneself with the equipment and procedures in this tutorial, recording physiological data can sometimes be challenging even for experienced researchers, and thus a fair amount of hands-on time will be necessary to train and supervise research assistants to ensure that high-quality data are collected (see Part III—> Experiment planning—> Training RAs). While conducting experiments, additional time ranging from 10 to 30 min per session must be budgeted for setting up and affixing equipment, and further allowances may be necessary for troubleshooting technical issues. Even relatively simple studies will likely require a minimum of hour-long sessions, in addition to extra time cleaning and preparing between participants. Post-collection, the resulting cardiovascular data need to be processed by a practiced researcher or research assistant, and the time required to manually process materials for each participant file (e.g., see Part IV—> Data cleaning and preprocessing—> Marking the B point; *Censoring waveforms*), should be considered in your planning.

Lastly, psychophysiological measures can be quite sensitive to the experimental environment. As a result, the recording environment needs to be relatively quiet and spacious enough to house the participant, hardware, computer, and potentially several other people (e.g., confederates) if required by your design, and it should maintain a comfortable temperature. It should ideally be located away from large sources of electrical interference such as HVAC units and elevator motors (Curtin et al., 2007).

Equipment and lab setup

The following sections provide an overview of the basic requirements of a psychophysiology lab and describe the proper use of equipment and software. The goal is to help researchers make informed decisions when outfitting their lab for the first time. Broadly speaking, a psychophysiology research setup consists of the following components:

Lab space – quiet, spacious, maintained at a comfortable temperature, and free from major sources of electrical interference. The room must be large enough to comfortably fit the participants, the researcher and any confederates, during both setup and data collection (if necessary), while ensuring the space is organized so as to avoid barriers or hazards (e.g., from wires). Since various devices can produce electromagnetic interference, it is a good idea to keep the physiological measuring equipment at least 18 inches away from computers (particularly power supplies and monitors), power cords, and speakers. The experiment should also be in a private location given the need for participants to lift their shirts to attach equipment.

During data collection, researchers may wish to remain nearby in order to monitor and intervene as needed to ensure smooth and continuous data collection. Should you prefer that the participant be left alone during the experiment, you may need to design creative solutions to create a second monitoring location, such as making partitions with privacy screens or setting up a remote control room. Either case requires a wired connection between the physiological equipment and a computer at the second location for data recording and real-time inspection of signals, as well as a means of controlling a stimulus computer if applicable (see below). If opting for a remote control room, you will also need a means of monitoring the participant with video and audio, and may also require an intercom system for communicating between the two rooms.

Computers – at least one stable computer with up-to-date specifications and sufficient storage for recording, storing, and analyzing physiological signals. Dedicating one computer to data collection is advised in order to continuously monitor incoming signals. Depending on your task, more computers may be necessary or desirable. For instance, your task design may require a dedicated computer for automated stimulus presentation, in addition to the one recording the data, or you may opt to use a laptop for portable data recording and conduct data analysis with a desktop workstation. If requiring participants to fill out surveys or do other behavioral tasks, an additional laptop or tablet may be necessary.

Hardware transducers and amplifiers – hardware units that measure and amplify physiological signals and convert them to a digital format readable by a computer. These systems are explained in detail in the following sections (Electrocardiography (ECG), Impedance cardiography (ICG), Blood Pressure (BP)).

Recording software – software that interfaces with amplifiers to record signals, often sold with hardware amplifiers (explained in detail in Choosing software below).

Analysis software – software used to process signals and extract variables of interest, often but not necessarily the same as recording software (explained in detail in Choosing software below).

Choosing hardware

While many standalone devices exist for recording each of the cardiovascular measures required here, the most convenient solution for outfitting your lab will likely be a modular research system. These systems, which consist of a central unit and removable modules, allow for simultaneous recording of the necessary signals and customization to suit the needs of your study. Additionally, new modules can be subsequently added to expand the capabilities of your system to meet future research needs.

At the time of writing, two such systems are best suited for acquiring the physiological signals implicated in challenge and threat research: the MP Research System from Biopac Systems Inc. (Goleta, California) and the Bionex System from Mindware Technologies LTD (Gahanna, Ohio). These systems are recommended because they offer modules for all three of the required techniques – electrocardiography, impedance cardiography, and non-invasive blood pressure. These systems also offer options for ambulatory measurements, as well as recording within a scanning (fMRI) environment. Other modular systems, such as the PowerLab system from ADInstruments (Dunedin, New Zealand), can also be made suitable with the addition of an impedance cardiography system (see Impedance cardiography (ICG) below) from another manufacturer, such as the VU-AMS (Vrije-Universiteit Amsterdam). Interested readers are directed to Jennings and Gianaros (2007) for further information about choosing a hardware vendor. Regardless of the research system chosen, the basic operating principles will remain the same.

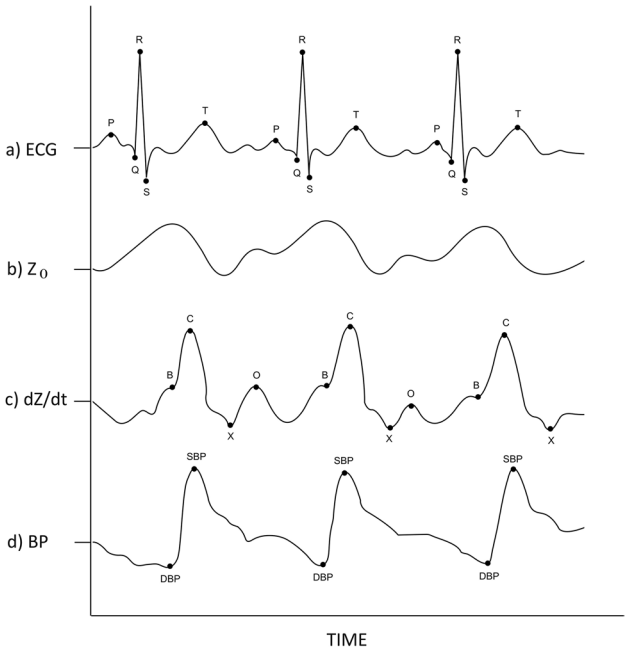


Fig. 1 Sample raw waveforms (ECG, Z_0, dZ/dt, BP)

Biosignals and filters

It is important to understand some fundamental concepts regarding the physiological signals you will be collecting and analyzing during your challenge and threat study. In essence, psychophysiological research entails capturing, analyzing, and interpreting biosignals—fluctuations of energy produced by a physiological process over time (e.g., changes in an electromagnetic field caused by contraction of the heart; Semmlow, 2018). Capturing biosignals requires a device capable of detecting and transducing these energy fluctuations into an analog electrical signal. This analog signal is sampled many times per second (the sampling rate) and the resulting values produce a digital time series, which is a list of values stored in chronological order that can be saved and manipulated by a computer. Plotting this series of values against time produces a waveform, such as the familiar electrocardiography (ECG) waveforms depicted in Fig. 1a.

When capturing and analyzing these time series, filters are commonly applied in order to attenuate certain problematic or irrelevant components of the signal and improve its interpretability. In order to understand how filters work, it is useful to understand that any periodic (repeating) signal can be represented as the sum of a number of simple sine waves of different frequencies and amplitudes (Fourier's theorem). The frequency of a periodic signal is simply the number of times it repeats over a given time interval, typically expressed in Hz (repetitions per second). Figure 2 demonstrates how several waves of different frequencies sum to form a more complex waveform. As such, applying a filter can be thought of as reducing the contribution of a subset of these component waves to the overall waveform. Depending on their specific parameters, filters can be designed to attenuate a wide range of components (e.g., all frequencies above 60 Hz) or a very limited range (e.g., a narrow band of frequencies centered on the power line frequency).

Filtering is a powerful means of increasing the signal-to-noise ratio of a recording, but can cause significant distortions and loss of information when used improperly (Luo & Johnston, 2010). It is therefore important to understand the difference between analog (online) filters applied at the recording stage, and digital (offline) filters applied at the analysis stage. At the recording stage, analog filters embedded in your recording hardware are useful for removing unwanted frequency components at the source, before the analog signal is converted to a digital signal. While useful, this means that hardware filtering should be applied carefully, as problems introduced at this stage cannot be fixed later. Digital filters are discussed in Part IV—> Data cleaning and preprocessing—> Digital filtering.

The basic filter types (applicable to both analog and digital filters) are described below:

Low-pass (LP): Attenuates frequencies above a cutoff point. For instance, a low-pass filter at 60 Hz means that frequencies lower than 60 Hz are allowed to pass through (hence "low-pass"), while frequencies above this value are attenuated. Figure 3a depicts a digital low-pass filter applied to an ECG signal contaminated with simulated high-frequency noise.

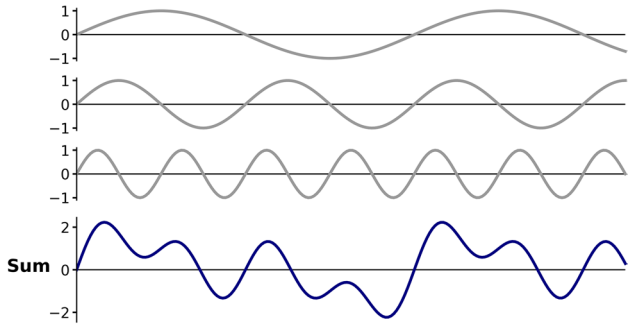


Fig. 2 A complex waveform (*blue*, *at bottom*) can be conceived of as the sum of a number of simple periodic (*sinusoidal*) waveforms (*grey*)

High-pass (HP): Attenuates frequencies below a cutoff, while allowing frequencies above to pass through. Figure 3b depicts a digital high-pass filter applied to an ECG signal with simulated baseline drift.

Band pass: Attenuates frequencies above and below a specified range, acting like a high-pass and low-pass filter put together.

Band stop: Attenuates frequencies within a specified range, allowing frequencies above and below to pass through. Band stop filters which attenuate a narrow range of frequencies around a target frequency (e.g., 60 Hz) are known as *notch filters*.

Electrocardiography (ECG)

Electrocardiography (ECG or EKG) is used to record the electrical activity of the heart using electrodes placed on the skin. These electrodes detect small changes in electrical potential at the skin caused by muscular activity (typically in the range of millivolts, mV), which are subsequently amplified and recorded as a digital time series. By placing these electrodes in an arrangement which triangulates the heart, the activity of the cardiac muscles (producing the prototypical ECG waveform; Fig. 1a) can be detected. It is important to realize, however, that ECG detects activity from all kinds of muscular activity, not just the heart, and is therefore susceptible to contamination by participant movement ("movement artifacts").

The record of cardiac activity over time that ECG provides is necessary for deriving heart rate (HR) and pre-ejection period (PEP) – two measures necessary for investigating challenge and threat. Heart rate (the frequency at which the heart beats) can be extracted from the ECG signal alone, while pre-ejection period (the time between electrical activation and contraction of the ventricles) requires both the ECG and ICG signal (see Part IV—> Data cleaning and preprocessing—> Deriving level I measures from raw waveforms).

ECG setup

Any ECG system will consist of three components: a hardware amplifier, electrode wires, and the electrodes themselves. By measuring differences in electrical potential between pairs of these electrodes, it is possible to record the electrical activity of the heart in a variety of different orientations, known as "lead configurations." For the current application, only a single lead is required (Blascovich et al., 2011), meaning that only three electrodes will be used – positive, negative, and ground (when using ICG concurrently, the ground electrode can sometimes be omitted; see documentation for your equipment for more details.) While other configurations are possible, the "Lead II" configuration is commonly used to study challenge and threat (e.g. Mendes et al., 2001; Seery et al., 2010; Tomaka et al., 1993). The standard Lead II configuration consists of a negative electrode on the right arm, positive on the left leg, and ground

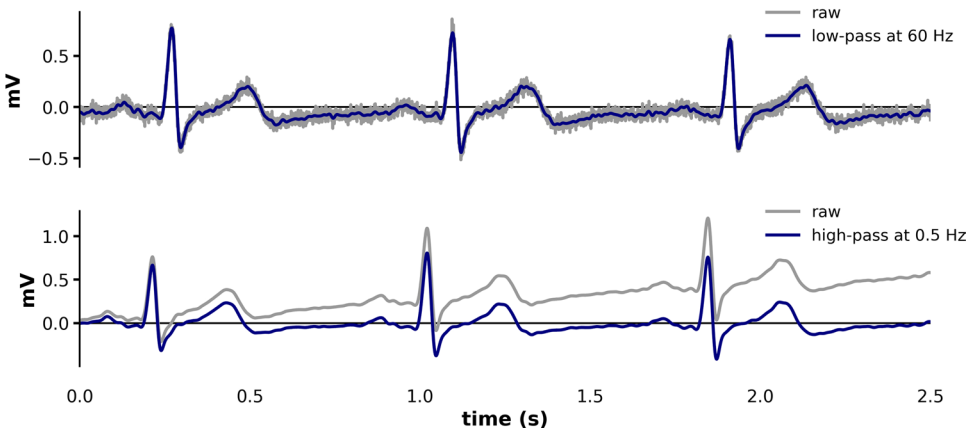


Fig. 3 Examples of digital filters applied to sample ECG data. Note: Zero-phase Butterworth filters (two-pass forward and reverse) were applied here offline for illustrative purposes. Effectiveness and amount of distortion will vary with filter type and design

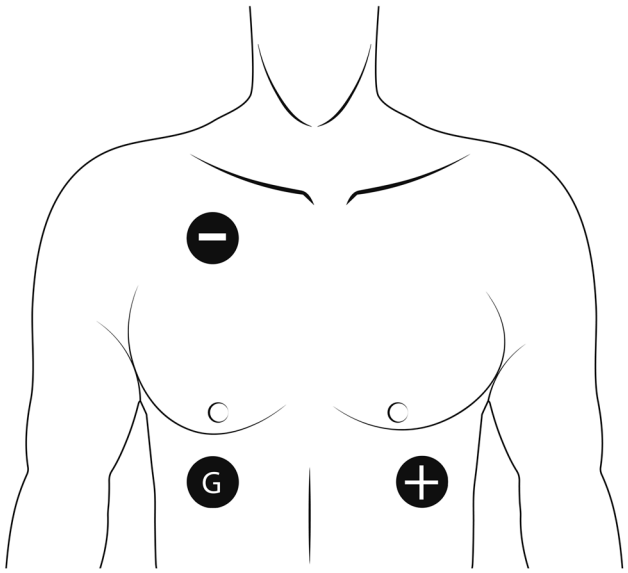


Fig. 4 Modified lead II ECG placement

on the right leg. Placement on the limbs, however, can result in motion artifacts, particularly when participants are highly engaged. As such, it is recommended that a modified placement be used, with the electrodes on the torso rather than the limbs: negative just below the right collarbone, positive just below the lowest left rib, and ground (if necessary) just below the lowest right rib (Fig. 4).

ECG hardware parameters

While the particulars of your ECG unit will vary, understanding the following general parameters will allow you to use most systems effectively. Take the time to understand these parameters, test them, and decide on appropriate values for your purpose at the outset. Once set, these parameters should be kept consistent for all participants.

Gain: The main function of the ECG hardware is to amplify the small electrical potentials measured at the skin before digitization. As such, an amplifier's gain setting reflects the ratio between the output and input voltages. For example, a gain setting of 1000 (a commonly used value) means the input signal is amplified by a factor of 1000. If the gain is set too high, the signal may exceed the range that can be digitally recorded, leading to "clipped" (flattened) peaks and therefore unusable data. Collect a small amount of test data prior to acquisition to ensure your gain is set appropriately and no clipping will occur.

Hardware filters: For ECG recording, a consortium of experts (Kligfield et al., 2007) recommends applying a hardware high-pass filter with a cutoff of 0.05 Hz and a low-pass filter with a cutoff of 150 Hz. This broad range preserves the necessary information in the ECG waveform, while excluding problematic signal components outside this range. In particular, the low-pass filter at 150 Hz prevents aliasing, a phenomenon where high frequency analog components (higher than the Nyquist frequency, i.e., half of the sampling rate) produce false lower frequency components during digitization. Additionally, the high-pass filter at 0.05 Hz eliminates any displacement of the average signal amplitude from zero, known as DC offset.

Impedance cardiography (ICG)

Impedance cardiography (ICG, sometimes ZCG or ZKG) is a non-invasive technique for estimating blood flow in the chest as a means of assessing cardiac output (CO) – a measure of cardiac efficiency (Sherwood et al., 1990), and one of the main indices of the challenge and threat model (Blascovich, 2008; Blascovich & Tomaka, 1996). ICG achieves this by exploiting the relationship between voltage and impedance in an electrical circuit, as expressed by Ohm's law:

$$V = I * Z$$

where V is voltage (the difference in electric potential between two points), I is current (the rate at which electric charge flows through the circuit), and Z is impedance (the resistance to current flow in an alternating current circuit). In a circuit where current (I) is held constant, voltage (V) varies directly with impedance (Z). By introducing a small alternating current of constant magnitude into the chest, an ICG system creates such a circuit in the body. Because blood is a good conductor of electricity, each heartbeat momentarily reduces the resistance to electrical flow across the chest, resulting in an increase in voltage as detected by a separate set of measurement electrodes. With this record of impedance changes, both the volume of blood pumped during a given heartbeat (stroke volume; SV) and the pre-ejection period (PEP) can be estimated (see Part IV—> Data cleaning and preprocessing—> Deriving level I measures from raw waveforms). It is worth emphasizing here that these values are *estimates*, in that they rely on several assumptions and approximations. For this reason, absolute values derived from ICG, especially volume-based measures like SV, and its derivative cardiac output (CO), have limited interpretability on their own. This is why measures of *change* (reactivity scores) are typically considered rather than absolute values (see Part IV—> Preliminary analyses—> Reactivity scores).

Impedance in the chest can also change for reasons unrelated to heart activity, including participant movement and the slow rise and fall of respiration. The latter produces a low-frequency signal (i.e., less than 0.5 Hz) which is often considered an artifact and removed, but does carry information of psychophysiological interest and may therefore be worth maintaining (Ernst et al., 1999). While beyond the scope of this tutorial, some authors choose to use respiration

as a covariate, or combine it with ECG to derive respiratory sinus arrhythmia, which indexes parasympathetic control (Cacioppo et al., 2007).

ICG setup

Similar to ECG, an ICG system will consist of a hardware amplifier, electrodes, and the wires connecting them. Unlike ECG, however, ICG involves an outer pair of current source electrodes in addition to the inner pair of voltage measurement electrodes. Three common variants of this basic configuration will be outlined here (Fig. 5).

The first decision to be made is whether to use band electrodes or an array of spot electrodes.[5] Band electrodes are conductive strips that typically encircle the participant's body, whereas spot electrodes are small, circular conductors attached to the skin by an adhesive patch (often ECG or similar electrodes are used). The four Mylar band electrode configuration, with two pairs of bands at the neck and torso, is standard (e.g., Blascovich et al., 2001; Mendes et al., 2001). Spot electrode arrays are also commonly used, and are sometimes preferred for their easier application, lower cost, and minimal discomfort compared with band electrodes. It is worth noting, however, that signal quality is often better with band electrodes, and less uniform current distribution with spot electrodes may make their measurements more susceptible to individual factors such as body shape (Woltjer et al., 1995). In general, challenge and threat research can be conducted successfully with either type, as long as planned comparisons involve change (reactivity) scores of cardiac output and stroke volume (Llabre et al., 1991; Tomaka et al., 1993; see Part IV—> Preliminary analyses—> Reactivity scores) rather than absolute values, especially if electrodes are to be reapplied between recordings.

If using band electrodes, a pair of bands separated by at least 3 cm will be placed around the neck, and another around the torso, at the level of the xiphisternal junction (where the lowest ribs meet, at the bottom of the sternum; Fig. 5a). For further details about this configuration, see the impedance cardiography guidelines published in Sherwood et al. (1990).

If using spot electrodes, there are two electrode configurations commonly used. One popular configuration is a minimal four-electrode array, developed by Qu et al. (1986). This array places a current electrode on the back of the neck and midway down the spine, and a measurement electrode on the front of the neck and in the middle of the chest (Fig. 5b). In this way, the function of each band electrode is approximated by a single spot electrode.

A variant of this configuration uses two electrodes to replace each band electrode, resulting in eight electrodes in total. A pair of electrodes (one current, one measurement) are placed on both sides of the neck and chest (Fig. 5c). Eight individual electrodes can be used, but paired electrodes are also available for this purpose, and are preferable for keeping a fixed distance between current and measurement electrodes. If using this array, ensure that electrodes are placed symmetrically.

The needs of your research project, the options available from your chosen manufacturer, and ultimately your preference will determine which of the three configurations outlined above should be chosen.

Note that dead skin may increase resistance to the electrodes, therefore cleaning or scrubbing the attachment site may improve connection in the case of ICG spot electrodes. This is less of a concern with band ICG electrodes, which cover a larger surface area. Regardless of the configuration chosen, once setup is complete, it is important to record the distance between the two measurement electrodes (or set of electrodes), as well as the participant's height, weight, and sex. These measurements are key to the accurate estimation of stroke volume and thus cardiac output.

ICG hardware parameters

As with ECG, ICG hardware parameters should be decided at the outset and kept consistent for all participants. Note that not all of the parameters below will be adjustable on some equipment (refer to the documentation for your device).

Range/Scaling: This parameter sets the ratio between the measured voltage and the reported impedance magnitude. Refer to the documentation for your hardware when setting this value, and collect a small amount of test data prior to acquisition to ensure your range/scaling is set appropriately to avoid clipping.

Frequency: The frequency of the constant alternating current (AC) passed through the chest. Typical values are 50 kHz and 100 kHz, and your choice should follow the manufacturer's recommendation for your equipment.

Filters: While there are no published guidelines on appropriate hardware filter bandwidths for ICG, similar principles apply as with EEG. Namely, hardware filters should be applied conservatively, allowing a broad range of frequencies to pass through (Hurwitz et al., 1993) while still preventing aliasing and DC offset. Filtering options will vary with hardware, but these requirements should be

[5] Amplitude-based measures (e.g., SV, CO, dZ/dt max) are highly correlated between spot and band electrodes, but differ in absolute value. Time-based impedance measures (e.g., PEP, LVET) are very consistent across measurement techniques. Thus, while absolute values cannot be compared between spot and band electrodes, percent changes and time-based measures can be (Gotshall & Sexson, 1994; Sherwood et al., 1992).

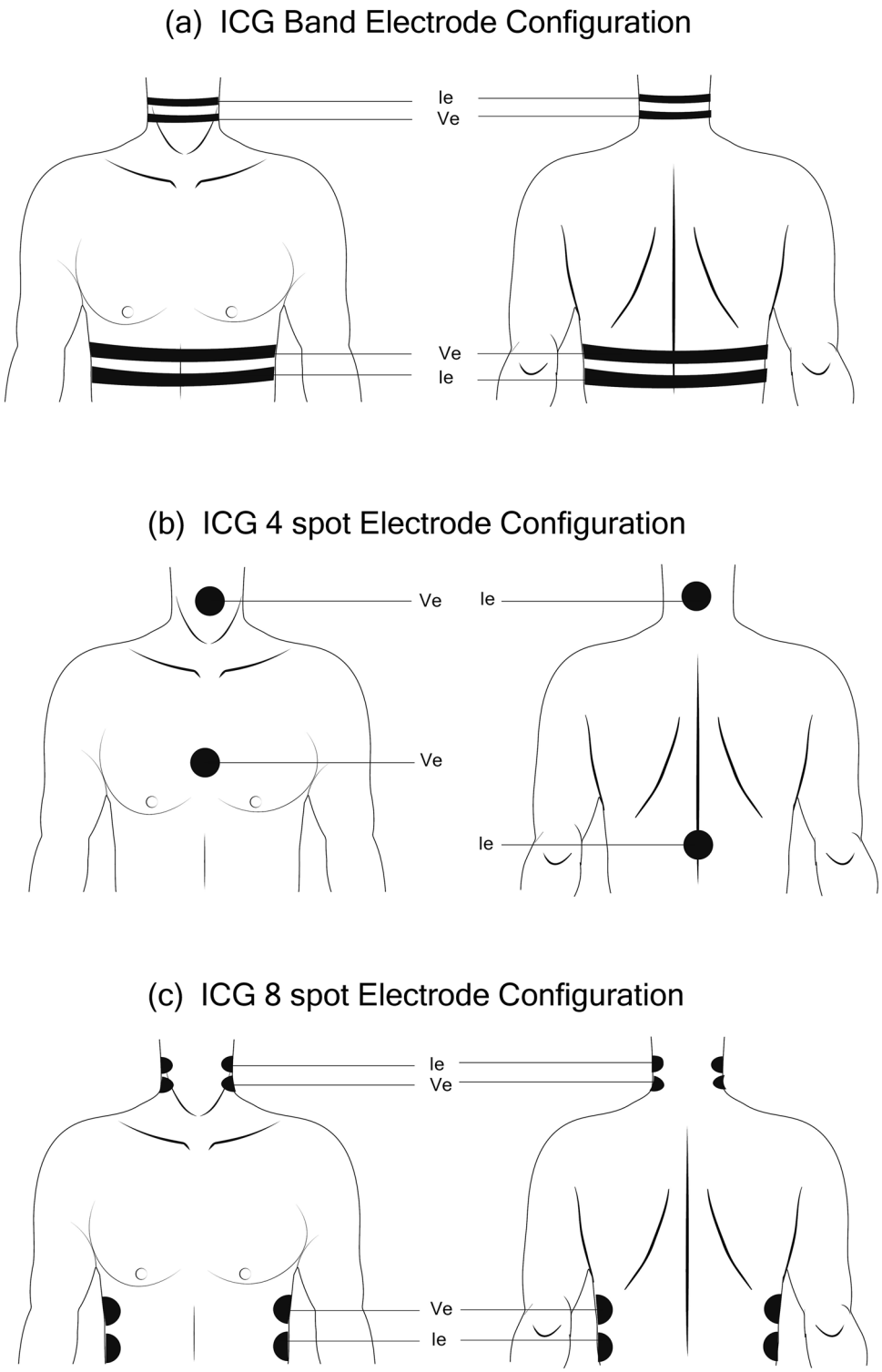


Fig. 5 Common ICG electrode configurations

met by a high-pass filter with a cutoff at 0.05 Hz or lower, and a low pass filter with a cutoff between 100 Hz and the Nyquist frequency of your data (one half the sampling rate).

Blood Pressure (BP)

In general, blood pressure (BP) measurement involves quantifying the force exerted by the blood on the walls of a

major blood vessel. This force varies over the heart cycle, where systolic blood pressure (SBP) is the highest value, and diastolic blood pressure (DBP) is the lowest. Mean arterial pressure (MAP) is a weighted average of SBP and DBP, and is a key determinant of total peripheral resistance (see Part IV—> Data cleaning and preprocessing—> Calculating level II measures).

Blood pressure can be measured non-invasively using a variety of techniques, which differ in their reliability, temporal resolution, and cost (see Blascovich et al., 2011, and Cacioppo et al., 2007, for further discussion). While the specifics vary with manufacturer and model, blood pressure systems are generally sold as standalone units, but many can output an arterial waveform or discrete SBP and DBP values to an external device, and can therefore be interfaced with the modular research systems described above. Broadly, these systems fall into two categories: intermittent and continuous (Meidert & Saugel, 2018).

Intermittent BP measurement techniques generally employ an inflatable cuff to occlude the brachial artery (a major blood vessel in the upper arm) and determine blood pressure from the resulting perturbations of blood flow. While this is most accurately performed by a trained professional, the oscillometric technique offers a practical, automated alternative, which determines blood pressure algorithmically. These systems can generally be configured to record at pre-set intervals, but the minimum length of these intervals varies by device. Given that challenge and threat data are generally analyzed in 1 or 2-min windows, an oscillometric machine that can acquire multiple BP readings within that interval is desirable.

Continuous, beat-to-beat blood pressure measurement can be achieved with tonometric or volume clamp technology. Tonometric methods place a sensor over the radial artery at the wrist, where the arterial waveform can be registered (sometimes with an additional reference cuff at the brachial artery in the upper arm). While this method produces frequent readings, it is sensitive to motion artifacts, and extra steps should be taken to immobilize the participant's hand and arm as much as possible (Blascovich et al., 2011). Also, blood pressure readings will vary inversely with vertical position relative to the heart, so the participant's hand should be kept as close to the level of the heart as possible, and this position should be maintained as closely as possible throughout the study. Manufacturers recommend using a sling or cradle to immobilize and correctly position the hand, if allowed by experimental design (Blascovich et al., 2011).

The volume clamp (or Peñaz) method can also provide beat-to-beat blood pressure readings using a combination of finger clamp and photodiode (as well as a brachial reference cuff in some cases). Generally speaking, the photodiode measures blood flow in the finger, while the clamp applies counter-pressure to keep blood flow constant. The pressure required to keep blood flow constant corresponds to the arterial pressure. This method is also sensitive to movement, so efforts should be made to immobilize the participant as much as your design allows.

Due to the range of measurement options available, as well as their relative advantages and limitations, choosing a blood pressure system can be difficult. In making this decision, it is helpful to note a few tradeoffs. Ideally, you want to record as many measurements as possible during the interval of interest, but the cost and complexity of the machine will increase substantially with higher measurement frequencies. Machines with higher sampling frequencies also generally measure from a location farther down the arm, making them more prone to motion interference. Furthermore, continuous measurements require additional steps to pre-process, and may drift and require recalibration at inopportune times. If capturing fast changes in BP and TPR is not a major priority, researchers may therefore opt for intermittent BP measurements, which are sufficient in most cases.

No matter which technique is used, participant comfort is an important consideration, especially for longer studies. Repeated blood pressure measurements with oscillometric and tonometric machines that contract frequently can cause discomfort and distraction, potentially resulting in artificially raised blood pressure over time.

Finally, given the proprietary nature of the algorithms used in many blood pressure devices, it can be difficult to objectively assess their validity. To mitigate these concerns, before purchasing a blood pressure system, be sure to verify that it meets the standards of the Association for the Advancement of Medical Instrumentation or the British Hypertension Society (Cacioppo et al., 2007), and be sure to report the model of device used in any subsequent publications.

Choosing software

In addition to the hardware discussed above, you will also need software to record and analyze your data, as well as an up-to-date and stable computer which meets the specifications for your software of choice (available in the software documentation). When purchasing from a well-established manufacturer, the most convenient option for acquisition and analysis will be to use the commercial software sold alongside your choice of hardware. This decision comes with the benefit of customer support provided by the manufacturer, as well as a large active user base. However, using commercial software comes with the downside of cost, as well as reliance on proprietary algorithms which can be opaque, inflexible, and can limit reproducibility.

Alternatively, it is possible to acquire data with one software package and analyze it with another, which allows one to take advantage of the (currently limited) open-source

options for physiological data analysis. In addition to being free, these open-source tools can offer greater transparency in analysis (in line with open science principles) and in some cases finer control over analysis parameters.

At the time of writing, the options below are recommended for their overall suitability for challenge and threat research, and for requiring the least amount of expertise and/or customization, although other options do exist (e.g., ADInstruments' LabChart, *pyphysio*).

AcqKnowledge is the comprehensive data acquisition and analysis package sold with Biopac's MP line of modular research systems. It can compute all of the necessary indices for challenge and threat research within one software suite, and supports ensemble averaging—a method of averaging signals across heart beats to create a single, representative waveform (see Part IV—> Data cleaning and preprocessing—> Ensemble averaging).

BioLab is Mindware's data acquisition platform, sold with the BioNex modular research system. For analysis, Mindware's Impedance Cardiography (IMP) Analysis Software can be used to analyze ECG and ICG data with ensemble averaging, and the Blood Pressure Variability (BPV) package can be used to analyze BP data.

MEAP (Moving Ensemble Analysis Pipeline; Cieslak et al., 2018) is an open-source, standalone application for analyzing ECG, ICG and continuous BP data using ensemble averaging. MEAP can import data acquired in any acquisition software, and allows close inspection and censoring of the ICG data in order to minimize the effect of motion artifacts, as well as manual control over placement of key waveform features (see Appendix B for more information on censoring waveforms; Cieslak et al., 2018 for further explanation and validation).

Open ANSLAB (Autonomic Nervous System Laboratory; Blechert et al., 2016) is an open-source MATLAB toolbox designed to compute a variety of physiological indices of ANS activity, including those necessary for challenge and threat research. Open ANSLAB also supports ensemble averaging and manual control over placement of waveform points. Note that, while Open ANSLAB itself is free and open-source, it requires a paid MATLAB license.

Acquisition software setup

After installing your software as per the manufacturer's instructions, the following areas will need to be considered before data collection can begin. Saving these settings as a template/configuration file for use with every participant will save time and ensure consistency across participants.

Channel mapping Each waveform must be assigned to a specific software channel. ICG outputs two signals: Z_0 and dZ/dt. Choose an order and color coding (if available) that aids your quick recognition of each signal.

Sampling rate Sampling rate is the number of times per second that an analog signal (e.g., voltage) is measured and recorded as a digital value. Higher rates offer better signal fidelity, with rates at or above 1000 Hz (1000 samples per second) being sufficient for this application (Blascovich et al., 2011).

Acquisition length If setting the acquisition length at the outset, ensure that it is set slightly longer than the expected length of your experiment to prevent data acquisition from terminating early. Researchers may elect to have one continuous acquisition file for the entire session, or start and stop measurements during relevant intervals (i.e., baseline, tasks, etc.). The former option may be preferable, however, since data storage is unlikely to be a limiting factor with modern hard drives, and making sure to start and stop the recordings at the appropriate times requires extra diligence on the part of the experimenter.

Event marking During data acquisition, you will need to mark the beginning and end of experimentally relevant time windows. This can be done manually, or using automated triggers from an external presentation computer or other stimulus presentation system. Refer to the documentation for your specific software to determine how to add time-stamped event markers to your file during recording. Give these event markers descriptive labels to facilitate interpretation.

Part III: Experiment planning and data collection

This section outlines important considerations for designing and executing your experiment, with an emphasis on practical issues related to physiological data collection.

Experiment planning

Sample size & other considerations

The standards for sample size in physiological research have evolved in recent years reflecting the greater emphasis on effect sizes and power analyses in psychological research more generally (e.g., Patil et al, 2016; Stanley et al, 2018). Whereas previously sample sizes were based on general rules of thumb like "20 participants per condition", increasingly researchers are turning to using power analyses to determine sufficient sample sizes a priori.

In order to conduct a power analysis one must first have an anticipated effect size. Unfortunately, the under-reporting of statistical procedures has undermined efforts to determine average effect sizes as well as reproducibility of cardiovascular physiology research more generally (Lindsey, et al., 2018). Estimates of effect sizes have ranged greatly, likely depending in part on the outcome of interest. A recent meta-analysis by Behnke and Kaczmarek (2018) examined the relation between challenge and threat indices and performance outcomes across 19 studies ($N = 1054$), finding small, but stable, average effects (*rs* for CO, TPR, and the combined Challenge Threat Index (CTI) range from |.10| to |.14|). Such evidence suggests that the conservative researcher should utilize small estimates of effect sizes in their power analysis. Considering effect sizes for self-report measures of challenge and threat, these have also ranged widely (Hase et al. 2019), likely due to the wide range of ways this is assessed (Behnke & Kaczmarek, 2018). As well, depending on the study design and nature of the appraisal, the researcher may or may not expect self-report and physiological measures of challenge and threat to align (e.g., in the context of interacting with an outgroup member; Mendes et al., 2002).

Statistical power, and thus sample size, depends also on the study design and planned analyses. Challenge and threat studies may examine both or either within- or between-subject comparisons. While within-subject, repeated measure designs provide more data points, increasing statistical power, they are not amenable to many research questions. As well, the researcher must consider issues such as the recovery period needed between stimuli and the potential for acclimation and order effects. Power analyses for within-person designs depend on the correlations between repeated-measures, which tend to be high. For example, in Behnke et al. (2020) the correlations from T1 to T2 are 0.52 for PEP, 0.87 for HR, and 0.77 for CO.

Another important consideration in determining sample size is the number of covariates that will be included in analyses. While many challenge and threat researchers use only baseline cardiovascular measures as control variables, others may wish to control for or examine the impact of other variables that may impact physiology such as age (e.g., Sillars & Davis, 2018) and BMI (e.g., Maier, et al., 2003) or variables that may impact task performance outcomes such as GPA, math anxiety, or even golf handicap (e.g., Chadha et al. 2019). Currently, most challenge and threat cardiovascular studies have sampled young, healthy, predominantly White undergraduate students. Thus, additional research is needed to determine whether, and to what extent, the distributions of relative challenge and threat responses vary for different populations (e.g., older vs. younger participants; Mendes, 2009).

Even with proper preparation, psychophysiological measurement is susceptible to equipment problems, artifacts, and experimenter error, which can necessitate the exclusion of some participants' data from analysis. Attrition from problems related to psychophysiological measurement varies widely between challenge and threat studies, but are typically in the range of 5 to 15% of participants recruited (e.g., Blascovich et al., 2004; Mendes et al., 2002; Mendes, Blascovich, et al., 2007; Vick et al., 2008), although values of 30% or more have been reported (e.g., Hoyt & Blascovich, 2010). It is therefore a good idea to account for an attrition rate of at least 10% (or more if your design is particularly lengthy or complex) when estimating your required sample size.

Minimizing movement

Although tasks that require speaking can generally yield usable data, subtle movements of the arms, legs, or chest introduce motion artifacts that can greatly reduce data quality. It is therefore crucial to minimize the amount of movement involved in all experimental tasks. Avoid moving the participant once the critical task has begun, and monitor them closely. If the experimenter is physically separated from the participant at any point (e.g., in a separate control room) consider using a video and audio monitoring system for this purpose. Verbal reminders to reduce movement should be given regularly, including before or during critical tasks. Opportunities for participants to rest and move in between critical tasks can reduce fidgeting. Soft physical constraints such as pillows or foam pads shaped to hold hands in the appropriate position (Blascovich et al., 2011) as well as psychological constraints such as signs that indicate where participants should place their hands or feet can also be effective at reducing movement and improving data quality. Note that arm movement may cause measurements to fail entirely in the case of some BP devices, so consider immobilizing the arm where BP measurements are taken (ideally with the sensor at or near heart height; see Part II—> Blood Pressure (BP)).

Baseline measure

Because levels of physiological arousal vary individually and are measured as a function of within-subject reactivity, it is important to include a baseline measure in your design. First, it has been recommended to let at least 10 min pass between attaching electrodes and commencing baseline measurements to allow ICG measurements to stabilize (Mohapatra, 1981). The most commonly used duration for the baseline period itself is 5 min, although longer durations can be employed to ensure that the participant has reached

a restful or baseline state of arousal. Typically, the last 1 or 2 min of this baseline period serves as the period of analysis, with the duration of the baseline analysis period matching that of the critical task (Frings et al., 2015; Tomaka et al., 1993).

Additionally, consider how the language of "baseline" may alert participants to upcoming tasks; calling it a "rest" or "calibration" period may therefore be preferable to reduce anticipatory effects. Given the impact of posture on physiological readings, it is important to ensure that the participant's physical position (e.g., sitting, standing, reclining) is consistent across the baseline and task periods.

Training RAs

Allot ample time to train and rehearse with research assistants. Each research assistant will need several hours to properly familiarize themselves and practice with the equipment, and multiple training sessions may be needed depending on the size of the team. This initial investment of time will ensure that they can competently and confidently run the experiment, which will yield higher-quality data, and help to instill trust in participants. Where possible, consider matching the gender of the experimenter to that of the participant or to the participant's preference to minimize concerns related to physical contact and partial undressing when applying the sensors.

Recruiting participants

To better inform participants, give them instructions in advance about the types of clothing to wear, such as discouraging one-piece outfits or tight-fitting clothes. When selecting participants, consider possible exclusion criteria early. For example, participants with hypertension, heart murmur, presence of a pacemaker, those taking cardiac medications, and those that are pregnant may need to be excluded (Jamieson et al., 2012) as well as anyone with a skin disorder or injury that would impede placement of the electrodes (e.g., severe psoriasis, new tattoos, sunburn). Depending on the nature of the task, researchers may consider additional exclusions or protections particularly where tasks are designed to induce significant levels of stress.

Data collection

Participant preparation

Once participants are settled, begin preparing the electrode attachment sites by wiping the skin at each site gently but firmly with a moist paper towel or alcohol swab. Apply a small bead of conductive gel to the electrode before attaching it firmly to the skin. Depending on the strength of your chosen electrode's adhesive, the climate where you are recording, and individual factors like hair, presence of dead skin, and proneness to sweating, you may wish to take additional steps to maintain proper contact throughout the recording session. For example, rubbing a small amount of abrasive gel (e.g., Nuprep Gel, Weaver and Company, Aurora, CO) on the skin may be necessary to remove dead skin and other debris. Care should be taken to avoid overabrading the skin, and ensure to wipe off the abrasive gel before applying the electrode. A strip of medical tape over the electrode may also be necessary.

After all electrodes have been attached according to your chosen configuration, connect the electrodes to the hardware modules, taking care to ensure they are connected correctly. Color coding systems and plans for how to direct wires away from the participant will help ensure correspondence and prevent disconnection or tangling.

Once the electrode attachment and configuration is complete, collect the following participant information necessary to calculate their stroke volume: sex, height, weight, and the vertical distance between the ICG voltage measurement electrodes (labelled *Ve* in Fig. 5). Although your chosen stroke volume equation (see Part IV—> Data cleaning and preprocessing—> *Deriving level I measures*—> Deriving level I measures from raw waveforms) may not require all of these measurements, it is in your best interest to take all measurements in case you change your analysis plan.

Throughout the experiment, prioritize participant comfort and well-being to reduce anxiety, ensure valid baseline measures, avoid disruptions during tasks, and prevent early termination. Researchers should make every effort to inform participants of next steps and possible discomforts, explain how equipment will be attached, and let the participant be in control wherever possible. For example, allow participants to lift their own clothing when necessary, or try to point out anatomical landmarks themselves to minimize feelings of intrusion. To promote participant well-being, maintain a competent and professional demeanor, respect physical boundaries, and give breaks. The experiment may need to terminate early if participants express severe discomfort, or if cardiovascular measures indicate that they have become unduly stressed (e.g., if systolic pressure exceeds 260 mmHg or diastolic pressure exceeds 115 mmHg; Greene et al., 2000).

Data acquisition

Check all signals before the start of the experiment, and regularly during data collection. For an example of proper waveforms, refer to Part IV—> Introduction to waveforms, and Fig. 1. Mindware Technologies LTD also provides a

useful guide outlining proper signals and troubleshooting recommendations (Morgan, 2017: https://support.mindwaretech.com; Training Guides—> Impedance Cardiography (ICG) Training Guide—> Improving Data Quality: General Guidelines). Some noise in the signal is unavoidable. However, dramatic changes in the signals, such as flatlining, sudden increases or decreases in amplitude, and large drifts suggest a poor connection or other malfunction, and should be corrected as soon as possible. Table 2 lists possible causes and corresponding troubleshooting options for common problems.

Equipment removal and storage

Upon completion of data acquisition, save your data immediately to avoid any interference from hardware removal. It is also a good idea to back up your data promptly and regularly in at least two additional locations. Disconnect all leads and place them securely away from the participant. Where possible, you may choose to allow participants to remove their own electrode pads to avoid discomfort. Provide a means of cleaning off the excess electrode gel.

When storing equipment, keep pre-gelled electrodes in an airtight, sealed bag to prevent drying out. Avoid excessive wear on the electrode cables by coiling or wrapping them neatly and loosely. Ensure that all equipment has a secure designated place to be stored.

Special considerations in data collection

Self-report measures

Self-report measures of challenge and threat, alternatively referred to as cognitive appraisals or evaluations,[6] are another facet of BPS-CT research sometimes used in tandem with physiological measures. Initial studies explored the relationship between the two, finding that self-report measures predicted cardiac reactivity and vascular resistance, reacted in line with physiological measures in response to varied instructions preceding a task, and were not affected by physical manipulations of physiological response patterns (Tomaka et al., 1993, 1997). Subsequent research has included self-report measures as an additional dependent variable alongside physiological ones (e.g. Jamieson et al., 2012; Mendes et al., 2001; Mendes, Mendes, Gray, et al., 2007), or as a manipulation check (e.g. Kassam et al., 2009). Such tools are intended to measure perceived demands versus perceived resources, which are typically captured through one or more questions or statements for each,

Table 2 Common problems in data acquisition, possible causes, and troubleshooting options

Common Problems	Possible Cause	Troubleshooting
- No signal - Noisy or irregular signal - Signal is not changing - Significant drift (note that some drift is normal, but marked or sudden change may suggest poor connection and may lead to loss of signal)	Scaling error	Manually re-scale or autoscale the window in the relevant channels until the main features of the signal are clearly visible
	Cable disconnection	Ensure that each cable is properly attached at the electrode, module, and connection (e.g., extension or splitter) sites
	Poor electrode connection or faulty electrode	Check to see that electrodes are not impeded by hair, scar tissue or other obstructions. Try pressing down gently but firmly on electrodes one by one to detect problem site; re-prep, re-attach and re-gel new electrode if necessary
	Improper cable connection	Trace all cables from electrode to hardware module to ensure proper connection sites. Color coding systems help
	Improper hardware settings	Check that all settings on the hardware and in the software are as designated. Saturation or clipping may be occurring. Adjustment of some settings may be necessary (e.g., gain, range, filtering)
	Calibration error	Re-calibrate machinery (e.g., blood pressure device) if possible
	Hardware malfunction	Reset machinery if safe and able to do so
	Broken equipment	Consult manufacturer if necessary. Replace wires, splitters, modules or other pieces of equipment

[6] See Part I—> The Biopsychosocial Model of Challenge and Threat (BPS-CT) and Mendes et al., (2001) for discussion on the merits of 'evaluation' versus 'appraisal'.

measured on a Likert scale (e.g. "how threatening do you expect the upcoming task to be?", "I am uncertain how I will perform." vs. "How able are you to cope?", "I have the abilities to perform well"). Questions are typically asked before and/or after the motivational task, but after instructions are delivered. Threat indices are generally calculated as a ratio of demands to resources, however these measures can be analyzed separately as well (Jamieson et al., 2012; Kassam et al., 2009; Mendes et al., 2001, Mendes, Gray, et al., 2007; Tomaka et al., 1993, 1997).

Self-reports can be a valuable complement to physiological measures, but also carry certain drawbacks, including effects due to social desirability, inaccessibility of nonconscious evaluation processes, and the limitations of retrospective and prospective evaluations (Weisbuch-Remington, et al., 2005; Seery et al., 2010). Further, asking participants to self-reflect may actually interrupt or alter their experiences (Seery et al., 2010; Vick et al., 2008).

Dyadic data collection

Researchers may be interested in assessing relative challenge and threat not just of one individual participant, but of two individuals simultaneously within a social interaction (e.g., Peters et al., 2014). Indeed, research on physiological synchrony and influence is growing (e.g., Helm et al, 2018; Thorson et al., 2018). How such a study is orchestrated will, of course, depend on the research question and the nature of the dyad (is this a couple in a long-term relationship, or two participants who have not yet met?). We discuss a few key considerations.

First, dyadic studies present some practical challenges that require considering space and logistics. Generally, baseline measures should be taken with the two individuals separated, unless the researchers have a theoretical reason to desire baseline measures to be taken within the social context in which the interaction is about to occur. Keeping participants separated for the baseline measures is particularly important if the two individuals are strangers to one another and the experimental design capitalizes on this unfamiliarity. Once baselines are established, the two members of the dyad need to be brought back together for the interaction, something which can prove practically difficult given the often stationary nature of physiological equipment and limited lab space. One possible solution is to utilize a retractable wall that can be removed to combine the testing rooms in which baseline measures are taken (e.g., Peters & Jamieson, 2016). Alternatively, the researcher may detach and reattach the leads, but leave the sensors in place as the participant moves between rooms. Alternatively, the researcher may opt for the "ambulatory" Biopac model (Bionomadix Wireless Wearable Physiology, Biopac Systems, Inc.) or other similar models, which allows the participant to move locations with all leads still attached. Note that while marketed as "ambulatory", this equipment is still sensitive to motion artifact and so should be considered ambulatory only between periods of data collection. Again, depending on the research questions, participants may be able to remain in separate testing rooms and interact over intercom or live videofeed.

Another important consideration is to ensure that the physiological signals measured for each participant can be mapped on to one another with precise timing. This can be achieved through using one data acquisition system (e.g., Biopac MP150) integrating data from different modules for each participant and sending it to the same computer for collection and visualization. Alternatively, two separate acquisition systems and recording computers may be used so long as they are synchronized to one another or utilize synchronized time stamps from a third computer (e.g., the one delivering the experimental stimuli). For more information on assessing physiological influence within dyads, see Thorson and colleagues (2018).

Collecting data outside the laboratory

With the growth of "wearables" (e.g., smart watches) that collect physiological data, researchers are increasingly able to move outside of the confines of the laboratory when assessing physiological indices. While wearables are increasingly effective at assessing heart rate and heart rate variability, in particular, even these are not well-validated, often proprietary, and highly susceptible to motion artifact (e.g., Ryan et al., 2019). Cardiovascular indices of challenge and threat require continuous (or near-continuous) measures of blood pressure and impedance cardiography (in addition to heart rate), which have not yet been integrated into commercially available or similar wearables. Therefore, while it is possible to collect challenge and threat data outside of the lab, an experimenter's presence will likely still be required. As well, the potential impact of motion artifact will need to be considered outside the laboratory just as it is within it.

Researchers interested in assessing challenge and threat outside the lab may, therefore, wish to consider salivary cortisol and alpha-amylase to index neuroendocrine and autonomic activation, respectively, instead of or alongside cardiovascular measures (see footnote 4). Although salivary measures do not have the temporal resolution of continuous cardiovascular measures, they are reliable indices of their associated physiological systems, are relatively non-intrusive to collect, and can be collected without the presence of an experimenter (Ali & Nater, 2020; Blascovich et al., 2011; Cacioppo et al., 2007; Kirschbaum & Hellhammer, 1989).

Part IV: Data processing and analysis

Introduction to waveforms

After acquisition, your data will consist of four elements: ECG, Z_0, dZ/dt, and BP waveforms. What these elements represent, as well as their key components, are outlined below.

ECG

The ECG waveform (Fig. 1a) represents the electrical activity of the heart over time, measured as voltage changes at the skin. The Q, R, and S points (collectively known as the QRS complex) indicate the electrical depolarization of the ventricles, which initiates the main pumping action of the heart. The R point in particular is generally used to identify a heartbeat for purposes of calculating HR (and sometimes PEP; see footnote 7), as well as synchronizing cycles for ensemble averaging (see Data cleaning and preprocessing—> Ensemble averaging below).

Z_0 (Basal impedance)

The Z_0 (basal impedance; Fig. 1b) waveform represents the total impedance (resistance to electrical flow) across the torso over time. It is used to estimate blood flow in the chest. It is important that ICG parameters are selected appropriately (see Part II—> Impedance cardiography—> 15), as basal impedance is used directly to calculate stroke volume (SV).

dZ/dt

The dZ/dt waveform (Fig. 1c) is the first derivative of the Z_0 waveform. As such, it represents the rate of change in impedance over time, and makes key inflection points more easily identifiable. On this waveform, the B point and X point represent the opening and closing of the aortic valve, respectively.

BP

The blood pressure waveform (Fig. 1d) depicts the arterial pressure at the site of measurement. The highest point is the systolic blood pressure (SBP). The lowest point is the diastolic blood pressure (DBP). Note that, depending on the system, SBP and DBP may be represented in two separate channels, or simply as discrete values recorded at intervals.

Data cleaning and preprocessing

Digital filtering

After collection of the signals above, you have the option to apply digital filters using software to further attenuate noise in your recordings. Digital filters come in the same basic categories as analog filters (i.e., low-pass, high-pass, band pass, and band stop; see Part II—> Biosignals and filters). However, a digital filter operates by applying mathematical operations to a digital time series, whereas an analog filter consists of hardware components that manipulate an analog electrical signal. As such, digital filters can be designed to have a wide array of characteristics, and your analysis software will likely provide numerous options. The details of digital filter design are beyond the scope of this tutorial, but interested readers can refer to Cook and Miller (1992) for an introduction to filter design tailored to psychophysiologists, as well as a helpful decision tree which will help you navigate the specific options available to you.

Ideally, applying digital filters at this stage can improve the signal-to-noise ratio of your data by removing problematic signal components. However, as with analog filters, digital filters also cause distortions, including changes to the timing and amplitude of key waveform points. For this reason, it is best to apply digital filters *sparingly* and *purposefully* (i.e., to correct a problem), rather than as a default step in the analysis process (Widmann et al., 2015), and always clearly report the motivation and specific filtering parameters used. Furthermore, always maintain a copy of the raw (unfiltered) data to compare against your filtered data, and only apply digital filters to continuous data, not epoched (segmented) data, or data that contain discontinuities.

A common application of digital filtering is the removal of AC line noise in ECG or ICG recordings, if visibly present, using a narrow band-stop (notch) filter centered at the main power line frequency where the data was recorded (50 or 60 Hz; see Luo & Johnston, 2010 for discussion of this practice).

Another common approach is to apply a low-pass filter to signals that appear generally noisy or "fuzzy" (Fig. 3a). It is important to note that low-pass filtering of this kind serves primarily cosmetic purposes (Widmann et al., 2015), and ensemble averaging (see Ensemble averaging below) will also serve to minimize high frequency noise (including power line noise) that is not synchronized to the heartbeat, and does not carry the same risk of introducing systematic distortions to the signal. If you do opt to use a high-pass filter in this way, ensure that the filter cutoff is set high enough to minimize attenuation of the informative part of

the cardiac signal (i.e., not lower than 50 Hz; Hurwitz et al., 1993).

Lastly, digital filters can be used to control "baseline wander", a low frequency artifact that causes large, slow changes in the average amplitude of the signal. Applying a Butterworth high-pass filter with a cutoff of 0.5 Hz can reduce baseline wander in ECG signals without substantially distorting waveforms (Fig. 3b; see Lenis et al., 2017 for demonstration and further discussion of high-pass filter design).

Segmenting data

At this stage, you will have one or more continuous data files for each participant. The total length of these files will depend on your experimental design, i.e., the number and length of the included task(s) and baseline(s). Only some segments of these data will correspond to time windows of interest (e.g., baseline or task) to be analyzed. Because cardiovascular responses tend to habituate over time (Kelsey et al., 1999, 2004), these analysis windows should be as short as possible while still including an adequate number of cardiac cycles for ensemble averaging (see Ensemble averaging below). As a rule of thumb, challenge and threat research typically uses analysis windows of 1 min (Mendes, Blascovich, et al., 2007; Tomaka et al., 1993) or 2 min (Blascovich et al., 2004; Frings et al., 2015). Shorter windows (e.g., 30 s) may be assessed so long as the data is of sufficient quality (e.g., minimal artifact) such that the ensemble averages reflect meaningful values. Additionally, researchers may elect to use a moving ensemble average (Cieslak et al., 2018) to examine rapid changes in event-related designs while still benefiting from the noise reduction capacity of ensembling.

Consult the documentation for your software of choice to determine how analysis windows can be specified. In some cases, you may be required to export each of these windows of interest as a standalone file for further analysis.

Ensemble averaging

Ensemble averaging is a method of averaging across multiple heart beats to produce one representative waveform for a given window of interest (Kelsey & Guethlein, 1990). By aligning the R peaks of multiple heart beats and taking their average, this method allows cardiovascular activity during a window of time to be analyzed as if it were a single beat. This method is preferred over alternatives, which involve either prohibitively time-consuming manual scoring of each beat individually or automated scoring methods that can be unreliable, especially with noisier data (Árbol et al., 2017).

Additionally, ensemble averaging has the benefit of minimizing the influence of any artifacts which are not synchronized to the R peak, such as participant motion or speech. The dZ/dt waveform is particularly susceptible to such artifacts (Kelsey & Guethlein, 1990).

However, as with any averaging procedure, ensemble averaging should be used with care, and the time windows analyzed should be chosen to be as homogeneous as possible. More concretely, researchers should not average over two or more time periods where cardiovascular responses are expected to differ substantially (e.g., parts of a rest period and a stress task; Kelsey & Guethlein, 1990). Because the resulting waveform would be an average of two qualitatively different cardiovascular states, it would not be a valid representation of the true cardiovascular activity during either time period.

AcqKnowledge, Mindware's Impedance Cardiography Analysis (IMP) package, and Open ANS Lab offer ensemble averaging tools. MEAP (Moving Ensemble Analysis Pipeline; Cieslak et al., 2018) is an open-source alternative that offers fine control over the ensemble averaging process.

Marking the B point

Accurate calculation of SV, CO, and PEP rely on correct identification of several key inflection points on the ECG and dZ/dt waveforms. The B point in particular requires special attention as it is the only key inflection point that is not a local maximum or minimum (see Fig. 1c). Moreover, the shape of the dZ/dt waveform can differ substantially between individuals, and efforts to design algorithms to automatically place the B point have had limited success (Árbol et al., 2017). As a result, visual inspection of the dZ/dt waveform is still recommended to ensure correct placement of the B point.

After ensemble averaging, your software will allow you to manually edit the placement of the B point and other inflection points. Correctly placing the B point is a skill that can be developed with practice. To that end, Árbol et al. (2017) offer an invaluable guide for correct placement (Fig. 6), covering the full range of waveforms you will encounter when scoring your data.

Censoring waveforms

Some data analysis programs will allow you to censor waveforms on a beat-by-beat basis, thereby excluding noisy or otherwise problematic cycles from the ensemble average. This can be useful for the dZ/dt waveform in particular, which is more susceptible to movement and other artifacts. If choosing to censor on a beat-by-beat basis, you may also wish to decide what percentage of cycles can be excluded before a participant's data are discarded. For instance, if more than 25–30% of the cycles within a given time window need to be censored, it suggests that the data may be too compromised by noise or other

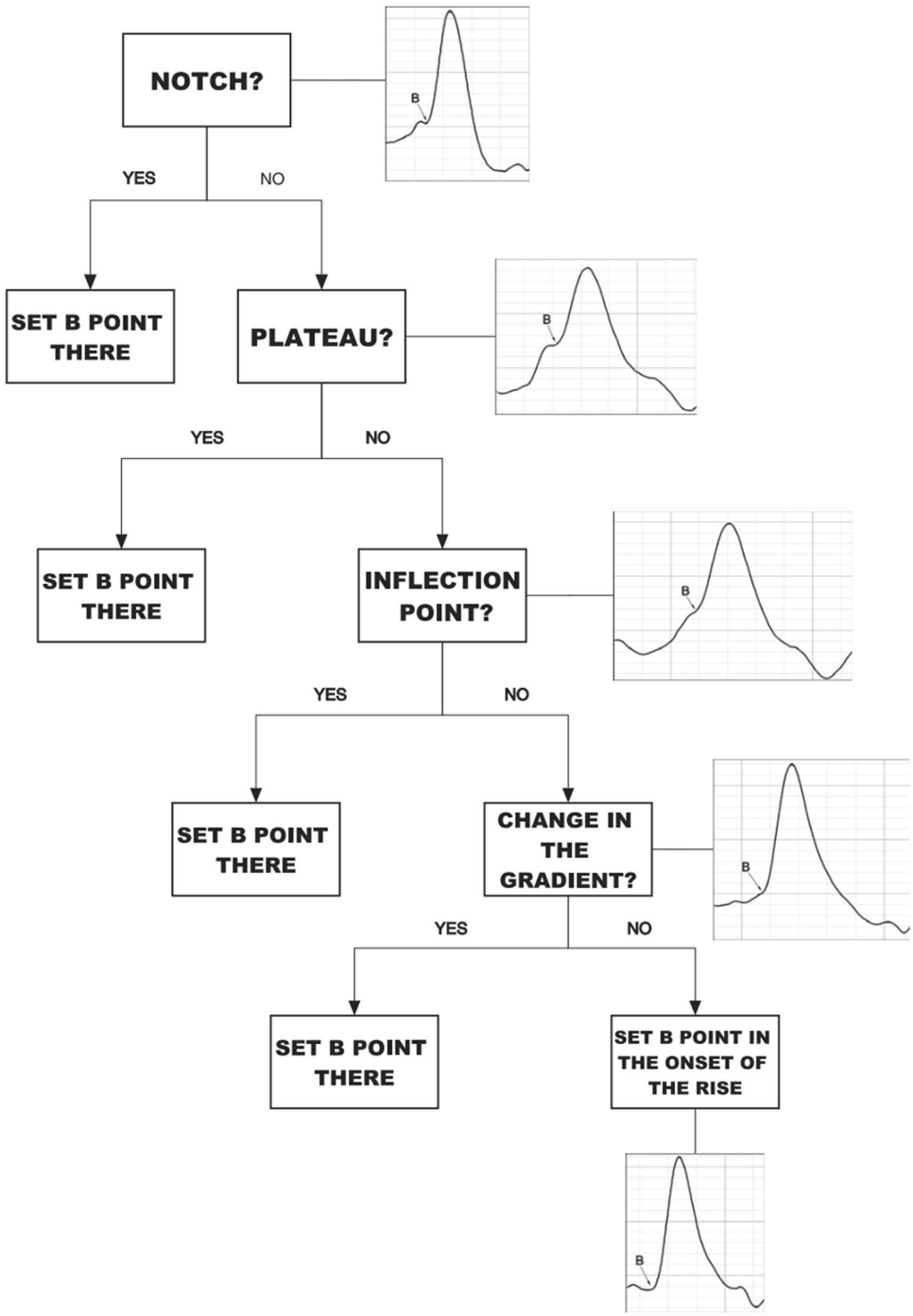


Fig. 6 Decision tree for B point placement on the dZ/dt waveform. *Note*: Reproduced with permission from "Mathematical detection of aortic valve opening (B point) in impedance cardiography: A comparison of three popular algorithms," by J. R. Árbol, P. Perakakis, A. Garrido, J.L. Mata, M. C. Fernández-Santaella, and J. Vila, 2017, *Psychophysiology*, *54*, p. 353 (10.1111/psyp.12799). Published by Wiley Periodicals Inc. Copyright 2016 by the Society for Psychophysiological Research

issues to give a valid measurement of cardiac activity, and should therefore be discarded entirely. This cutoff is up to the discretion of each researcher, but it is important that censoring criteria are set early and remain consistent for all participants. Appendix B provides a sample set of censoring criteria.

Table 3 Level I measures

Level I measure	Description	Calculation	Derived from
Heart rate (**HR**)	Number of heart beats in given interval	Number of R peaks (or QRS complexes) divided by length of interval	ECG
Pre-ejection period (**PEP**)	Measure of contractility – speed and force of heart contraction	Time (ms) between Q or R point (ECG; see footnote 7) and B point (dZ/dt)	ECG and ICG (dZ/dt)
Stroke volume (**SV**)	Volume of blood ejected from the heart in one beat	One of several equations (e.g., Kubicek, Sramek-Bernstein)	ICG and additional participant measurements
Mean arterial pressure (**MAP**)	Average pressure in arteries	$MAP = {}^{1}/_{3} \times ((2 \times DBP) + SBP)$	BP

Deriving level I measures from raw waveforms

For the purposes of this tutorial, "Level I" measures refer to those variables that can be derived directly from the physiological waveforms, whereas "Level II" measures are composites of two or more Level I measures.

Several Level I measures play important roles in indexing challenge and threat: heart rate (HR), pre-ejection period (PEP), stroke volume (SV), and mean arterial pressure (MAP). These are summarized in Table 3.

While software packages may differ somewhat in their derivation of these measures, the general principles remain the same and are explained here. After familiarizing yourself with the basics laid out below, refer to the software documentation to find the steps for calculating each measure in your particular software.

As a reminder, HR and PEP are rate-based (chronotropic) measures, which are considered valid time point estimates (Blascovich et al., 2011). On the other hand, due to limitations of the technology, variability introduced by sensor placement, and the approximative nature of the equations involved, SV, MAP, and the Level II measures derived from them (CO and TPR) should generally be considered only in terms of relative changes over time (reactivity scores), not in terms of absolute values (Blascovich et al., 2011). Although HR, SV, & MAP are included in the calculations of the primary outcomes CO and TPR, we recommend that each variable be reported individually as well. This allows for readers to ascertain the extent to which results are driven by the various components of each index of cardiovascular reactivity. For example, a significant increase in CO could be driven by increased HR, SV, or both. Condition differences driven solely by HR may be more indicative of differential SAM activation than relative activation of the HPA axis, which is the theoretical basis for the distinction between these two psychophysiological states (Blascovich et al., 2011; Seery, 2011). Indeed, early work examined HR (as well as PEP or VC) as potential indicators of challenge and threat, but these measures alone did not predict differential performance, but CO and TPR did (Seery, 2011).

Heart rate (HR) Heart rate is the frequency at which the heart beats, usually expressed in beats per minute (bpm). It is calculated by dividing the number of R peaks (or QRS complexes; see Fig. 1a) during an interval by the duration of the interval. It is therefore important to remember that heart rate is always the result of an averaging process, and the length of the averaging window should be considered carefully for your application. Similar to the above discussion (43; Ensemble averaging), averaging windows should be brief, represent periods where cardiovascular responses are expected to be consistent, and should not span across different tasks.

HR is required to calculate cardiac output (CO; see Calculating level II measures below). HR is often also used as an index of task engagement, but its dependence on interacting sympathetic and parasympathetic influences make it a less valid measure of engagement than PEP (Brownley et al., 2000; Schächinger et al., 2001; see Preliminary analyses—> Tests of Engagement below).

Pre-ejection period (PEP) Pre-ejection period is the time (in milliseconds, ms) between the electrical depolarization of the ventricles (marked by the Q point on the ECG waveform) and the opening of the aortic valve (marked by the B point on the dZ/dt waveform; see Fig. 7) when oxygenated blood enters the aorta on its way to general circulation (Sherwood et al., 1990).[7] As such, accurate PEP values rely on proper placement of the B point (see Part IV—> Data cleaning and preprocessing—> *Marking the B point*). Importantly, this interval is strongly influenced by sympathetic activity, and has been shown to vary with task demands (Blascovich & Tomaka, 1996; Kelsey et al., 2004; Wright & Kirby, 2001). As such, PEP is a more valid indicator of task engagement than HR (see Preliminary analyses—> Tests of Engagement below).

By convention, PEP reactivity is often recoded as ventricular contractility (VC), which is operationalized as the

[7] In practice, the interval between the R point and B point is often used instead, as the Q point can be difficult to mark, and the interval between Q and R tends to be constant (Berntson et al., 2004).

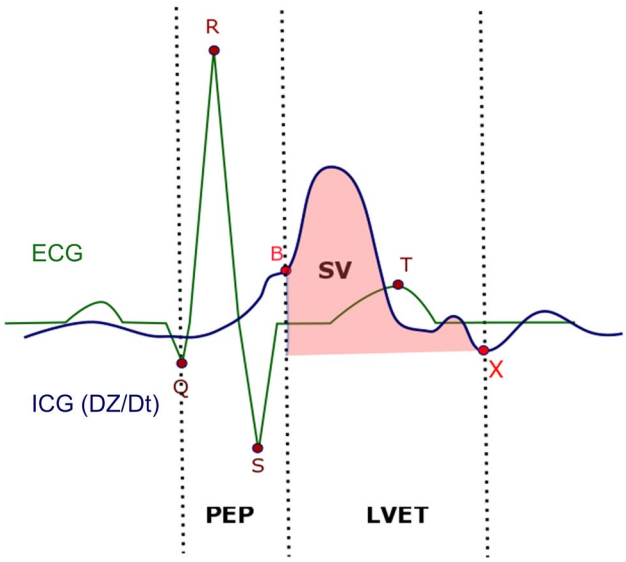


Fig. 7 Level I measures derived from the ECG (green) and dZ/dt (blue) waveforms

change in PEP between two timepoints multiplied by –1. As such, a positive VC value means a decrease in PEP, and indicates task engagement (Blascovich et al., 2004; Frings et al., 2012; Mendes et al., 2002).

Stroke volume (SV) Stroke volume is the estimated blood volume ejected from the heart during a single heartbeat. SV (along with HR) is a key determinant of cardiac output, where higher values are associated with the challenge state.

It is important to note that SV can only be estimated, using one of several possible equations, such as the Kubicek (Kubicek et al., 1970) or Sramek-Bernstein (Bernstein, 1986) equations. Each equation uses a different technique to estimate the fluid volume in the chest, based on participant measurements (e.g., distance between measurement electrodes [Kubicek], participant height and weight [Sramek-Bernstein]) as well as features of the Z_0 and dZ/dt waveforms. A key feature used by all SV equations is left ventricular ejection time (LVET). This interval, representing the time between the opening and closing of the aortic valve, is measured as the time between the B and X points on the dZ/dt waveform (Fig. 7). As with PEP, accurate SV values depend on proper placement of the B point (see Part IV—> Data cleaning and preprocessing—> *Marking the B point*).

Mean arterial pressure (MAP) Mean arterial pressure is a measure of average blood pressure during the cardiac cycle. MAP is a determinant of TPR, where higher values are associated with greater threat. MAP is calculated as a weighted average of systolic blood pressure (SBP; the point of highest pressure) and diastolic blood pressure (DBP; the point of lowest pressure), where $\text{MAP} = {}^{1}/_{3} \times ((2 \times \text{DBP}) + \text{SBP})$ (Blascovich, et al., 2011). This weighting reflects the difference in length between these two components of the cardiac cycle.

Calculating level II measures

After computing the Level I measures (HR, PEP, SV, and MAP) with your software of choice, calculating the Level II measures (CO and TPR) is straightforward. While your software may output CO and TPR along with the Level I measures, it is recommended that you calculate these values manually (see Part IV—> Dealing with outliers below).

Cardiac output (CO) Cardiac output is the volume of blood (L) pumped by the heart in one minute. It is considered a measure of cardiac efficiency. Higher values are associated with greater challenge and less threat (Blascovich & Tomaka, 1996; Blascovich et al., 2011). CO is simply calculated as the product of SV (ml) and HR (bpm), with a conversion factor to convert the unit to liters (Blascovich et al., 2011):

$$CO = \frac{SV \times HR}{1000}$$

Total peripheral resistance (TPR) Total peripheral resistance is a measure of the peripheral vasculature's overall resistance to blood flow. Higher values are associated with greater threat and less challenge (Blascovich & Tomaka, 1996; Blascovich et al., 2011). TPR is generally expressed in units of dyne-seconds · cm^{-5}, and is calculated as (Sherwood et al., 1990):

$$TPR = 80 \times \frac{MAP}{CO}$$

Dealing with outliers

After data cleaning and preprocessing, researchers must consider an approach to dealing with outliers. This includes identifying whether outliers are of enough magnitude to require modification (Reifman & Keyton, 2010). If so, the source of variation should be considered. While outliers from incorrect data entry, measurement error, syntax errors, or entries which are not a member of the intended population may require removal or correction, values from the intended population which show extreme values may require other treatment (Reifman & Keyton, 2010; Tabachnick & Fidell, 2007). Before engaging in methods to alter outliers, researchers need to decide whether the outliers at hand are

due to the errors mentioned above or are a true reflection of extreme values (Reifman & Keyton, 2010). Researchers may also consider whether variable transformation is a suitable method to retain normality (Tabachnick & Fidell, 2007).

If choosing to alter outliers, two options are trimming and winsorizing. Trimming refers to the method of simply removing a given number (k) of outliers on either end of the distribution. Winsorizing does not remove outliers, but transforms them to be less extreme. Outliers can be changed to the value of the closest non-outlying value or a set percentile (Reifman & Keyton, 2010). If trimming or winsorizing is a preferred method, a number of additional decisions must be considered, including whether or not to trim or winsorize the data symmetrically or asymmetrically (Reifman & Keyton, 2010), and what cut-off point to use for outlier selection. Example methods include a set standard deviation (Tabachnick & Fidell, 2007) and a fixed proportion of the sample, which may or may not take into consideration the actual pattern of data (Reifman & Keyton, 2010; Ruppert, 2006).

Although the choice is at the discretion of the researcher, winsorizing may be preferable in the case of small sample sizes in order to retain statistical power (Reifman & Keyton, 2010). In any case, ensure to report your decisions and justifications for your chosen approach, particularly if excluding outliers (De Veaux et al., 2012).

Both trimming and winsorizing can be performed on Level I and Level II measures. Therefore, it may be beneficial to calculate Level II measures manually from Level I measures after data-cleaning, rather than using software output of Level II measures. Critically, all treatment of outliers should take place prior to hypothesis testing.

Preliminary analyses

Reactivity scores

Challenge and threat researchers, as well as those using other psychophysiological frameworks, tend to use CO and TPR reactivity scores as dependent variables to examine the cardiovascular changes associated with psychological processes (Llabre et al., 1991; Tomaka et al., 1993). A reactivity score reflects the change in activation from baseline to task. It is calculated for each physiological measure of interest by subtracting the baseline value from the task value. This approach helps to overcome the aforementioned issues with interpreting raw values, which arise from the limitations of the measurement techniques and individual differences.

Tests of engagement

Challenge and threat states are theorized to occur in motivated performance situations, which require the participant to be actively engaged in the task at hand and be motivated to succeed (Blascovich & Mendes, 2000; Blascovich et al., 2011). Therefore, it needs to be demonstrated that participants are suitably engaged before interpreting data through the lens of the BPS-CT.

Two indices are typically considered for this purpose, often together. The first is PEP (sometimes recoded as ventricular contractility, VC), which has been validated as an index of sympathetic activation (Schächinger et al., 2001). Specifically, decreased PEP (increased VC) indicates greater sympathetic activation, and therefore greater task engagement under the challenge and threat framework. The second index is HR, which is a function of both sympathetic and parasympathetic control (Levy, 1977; Berntson et al., 1993). Under the view that engagement is proportional to the degree of sympathetic influence on the cardiovascular system (Obrist, 1981; Wright & Kirby, 2001), PEP may therefore be considered a more valid measure of task engagement than HR. While HR may be sensitive to sympathetic activity in situations of high demand, conclusions about task engagement should not be drawn from HR alone.

Challenge and threat index (CTI)

The Challenge and Threat Index (CTI) combines CO and TPR reactivity scores into one measure. This is achieved by first standardizing each participant's CO and TPR reactivity scores into Z-scores (Blascovich et al., 2004; Kassam et al., 2009). In other words, each participant should have one standardized CO reactivity score and one standardized TPR reactivity score for each baseline-task pair. Then the standardized TPR score is reverse-coded (multiplied by −1) and summed with the standardized CO score. This produces a single score for each baseline-task pair, where higher scores indicate more challenge-like reactivity and lower scores indicate more threat-like reactivity. The opposite pattern can be achieved by reverse-scoring CO instead, if desired for ease of interpretation or presentation purposes (Kassam et al., 2009). Either way, the CTI is a useful index for summarizing overall cardiovascular reactivity in a single variable. However, researchers vary in their choice of whether or not to combine CO and TPR into this overall index, and this choice depends on the goals of the research and the intended audience. While the CTI provides a clean and simple way to present results, authors should be clear about the particular components that drive these effects. Importantly, the CTI should not be used to obscure a lack of significance in either TPR or CO. Because TPR and CO are meaningful in and of themselves, we recommend that researchers report these values individually as well, at the very least as supplementary materials. Such comprehensive reporting is essential to increasing transparency in this area of research.

Primary analyses

Recall that under the BPS-CT, greater challenge is associated with an increase in CO and a decrease in TPR, whereas greater threat is associated with a decrease or no change in CO and an increase or no change in TPR (Blascovich & Mendes, 2000; see Table 1 for further details). On the basis of these patterns, one can test whether a manipulation produces cardiovascular responses suggesting greater challenge or greater threat compared to a control condition. Keep in mind, however, that challenge and threat are indicated by measures of reactivity or change, not absolute value. Therefore, different absolute values of cardiovascular measures between experimental and control conditions do *not* constitute evidence that the experimental manipulation has an effect. Instead, such a conclusion would require differences in the baseline-to-task *change* of cardiovascular measures between experimental and control conditions.

Suppose you hypothesize that a particular manipulation will increase participants' tendency to exhibit greater challenge. This hypothesis would be supported if the experimental condition (relative to the control condition) exhibited greater increases in CO and greater decreases in TPR from baseline to task. Note, however, that the Challenge and Threat Index (CTI) uses change scores in its calculation, so it is inherently a measure of reactivity and can be compared between experimental and control conditions. Unlike physiological reactivity scores, which measure change from baseline to task, self-report measures (see Part III—> Special considerations in data collection—> Self-report measures) represent subjective appraisals at one time point, and may be compared directly between conditions.

In creating statistical models, researchers may consider adding baseline measures as covariates or, if performing multiple trials of the same task, adding the initial task measures as covariates. This can be used to control for dependencies between baseline values of a measure and the magnitude of its response to a stressor (see discussions related to the "Law of initial values"; e.g., Stern et al., 2001; Berntson, Uchino, et al., 1994; Berntson, Cacioppo, et al., 1994). Such decisions are at the discretion of the researcher and should be accounted for in power analyses and included in the preregistered plan (if applicable) prior to data collection and analysis.

Transparency in reporting

By this point it will be clear that employing psychophysiological measures in your study adds considerable "researcher degrees of freedom", meaning that flexibility in data collection and analysis could allow different results and interpretations to be drawn from the same dataset (Simmons et al., 2011), potentially reducing confidence in the findings. This problem can be mitigated by preregistering your physiological processing pipeline along with your hypotheses and study design, but in practice pre-specifying every detail of an analysis pipeline may be unrealistic (e.g., the need for off-line filtering may only be apparent after data collection). Therefore, transparently and thoroughly reporting your data collection and analysis decisions is also necessary to ensure your work is interpretable and replicable by other researchers. To this end, we provide a checklist (Table 4) of necessary information to report when publishing your results, including often overlooked technical aspects of data collection and analysis. For further reading about open science practices in psychophysiology, we direct readers to Garrett-Ruffin and colleagues (2021).

Table 4 Methods reporting checklist

✓	Data collection and analysis decisions
	Manufacturer and model names of all acquisition equipment, including all modules (ECG, ICG, and BP)
	Blood pressure technique and sampling interval
	Sampling rate and other hardware parameters (including hardware filter settings, gain, ICG current frequency, and range/scaling where applicable)
	Electrode type used (e.g., band or spot)
	Electrode configuration used (for both ECG and ICG)
	Acquisition software used (including version number)
	Analysis software used (including version number)
	Equation used to calculate SV (e.g., Kubicek, Sramek-Bernstein)
	Methods used to calculate Level I measures (software-dependent)
	Digital filter parameters (filter type, cutoff frequency, filter order, transition bandwidth etc.; see Cook & Miller, 1992) and justification for their application
	Analysis windows (length and relationship to other events in the task)
	dZ/dt censoring procedure (if applicable)
	Missing data rates and thresholds for exclusion
	Winsorizing procedure and thresholds (if applicable)
	Any special considerations related to your particular research design (e.g., use of ambulatory equipment, dyadic recording etc.)

Concluding remarks

Research over the past 25 years has shown the BPS-CT to be a useful model that provides a validated link between physiological measures and psychological constructs. The model is flexible in its use, as the core requirements for eliciting challenge and threat (including goal relevance, evaluation, and active performance) can be easily incorporated into other methodological paradigms. Indeed, the occurrence and malleability of challenge and threat have

been demonstrated within standardized tasks such as the Trier Social Stress Test (Jamieson et al., 2012; Kassam et al., 2009), Remote Associates Test (Seery et al., 2009), and Graduate Record Examination (Vick et al., 2008). The validity and interpretability of this model, as well as its applicability across a variety of paradigms and fields of research, make it a powerful tool.

It is important, however, for researchers to be aware of the practical requirements of this tool. Although long-term payoffs can be rewarding, start-up costs can be high. While some systems (e.g., ICG, ECG) allow for precise measurements at reasonable costs, other systems (e.g., blood pressure machines) present a delicate balance between precision and price. Such decisions can be made by jointly considering the nature of the task, the norm in the relevant area of research, and the availability of resources.

Furthermore, the logic of psychophysiological inference is important, and should not be taken for granted. Using a validated model such as the BPS-CT increases confidence in drawing psychological inferences from physiological processes, but researchers should still be aware of the underlying logic and related challenges when interpreting results (see Part I—> What is psychophysiology and why is it useful?).

Overall, conducting a psychophysiological experiment based on the BPS-CT requires non-trivial resources and skills. By detailing the benefits, limitations, requirements, and practical steps above, this tutorial seeks to (1) inform researchers about whether and how to incorporate this tool into their methodological repertoire and (2) provide an outline for what kinds of methodological details should be routinely reported in research articles to enhance transparency and replicability. Our hope is that for researchers planning to conduct their first psychophysiological experiment, with the aid of this tutorial, the process will be less of a threat but more of a feasible and rewarding challenge.

Appendix A: Acronyms

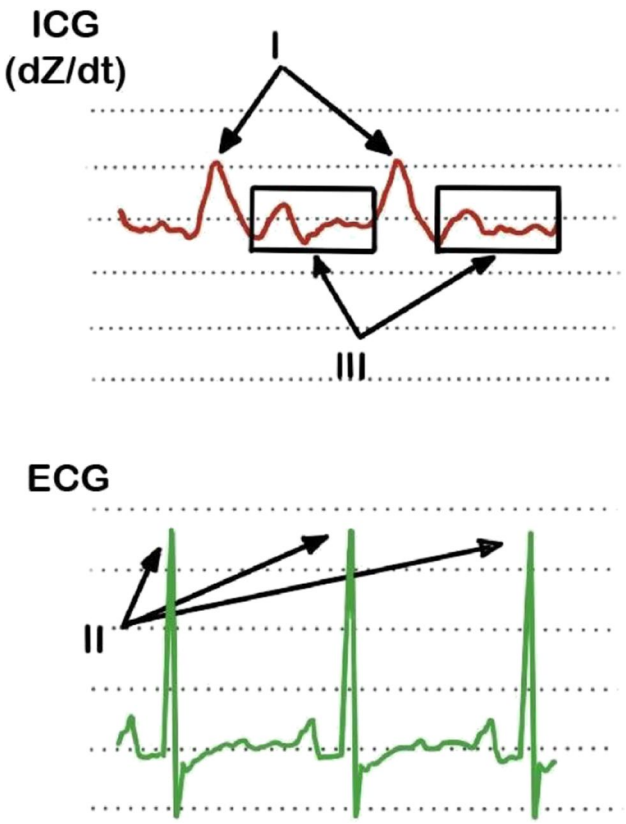
Acronym	Term
AC	Alternating current
ANS	Autonomic nervous system
BP	Blood pressure
BPM	Beats per minute
BPS-CT	Biopsychosocial model of challenge and threat
CO	Cardiac output
CTI	Challenge and threat index
DBP	Diastolic blood pressure
DC	Direct current
ECG or EKG	Electrocardiography
HP Filter	High-pass filter
HPA axis (or PAC axis)	Hypothalamic–pituitary–adrenal axis
HR	Heart rate
ICG or ZCG or ZKG	Impedance cardiography
LP Filter	Low-pass filter
LVET	Left ventricular ejection time
MAP	Mean arterial pressure
MEAP	Moving ensemble average pipeline
PEP	Pre-ejection period
SAM axis	Sympathetic–adrenal–medullary axis
SBP	Systolic blood pressure
SV	Stroke volume
TPR	Total peripheral resistance
VC	Ventricular contractility

Appendix B: Data censoring and point detection

This appendix provides examples of censoring criteria prior to ensemble averaging the ICG waveform, using MEAP's Inspect Data function (Cieslak et al., 2018). These examples are meant to serve as a guide of possible censoring criteria, based on the combined experience of the authors. Although censoring criteria may vary by researcher, it is important to remain consistent across participants within a particular investigation. We recommend establishing a firm cut-off for the amount of censored data after which the participant is excluded. For example, you may decide that having to censor > 25–30% of the cycles results in exclusion. For further recommendations regarding B point detection after ensemble averaging, see Figure 6, and VU University Ambulatory Monitoring System's (VU-AMS) resource 'Impedance scoring' (VU-AMS Ambulatory Monitoring System, n.d.: http://www.vu-ams.nl/support/previous-ams-versions/manuals/amsimp/impedance-scoring/).

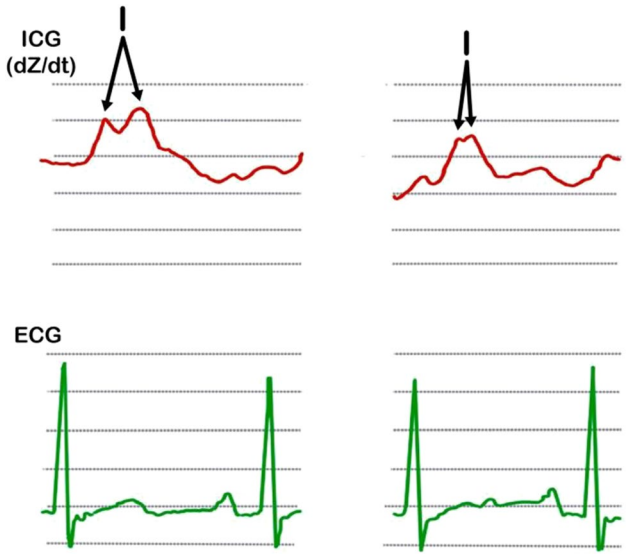
Main features of an ICG (dZ/dt) cycle:

Clear peak (I) immediately after the corresponding ECG peak (II)
Smaller peaks/variation (III) following the first peak

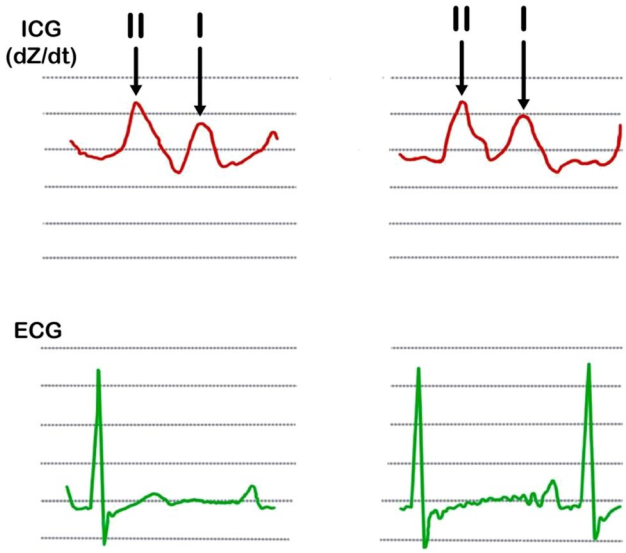


Minor deviations (no censoring recommended):

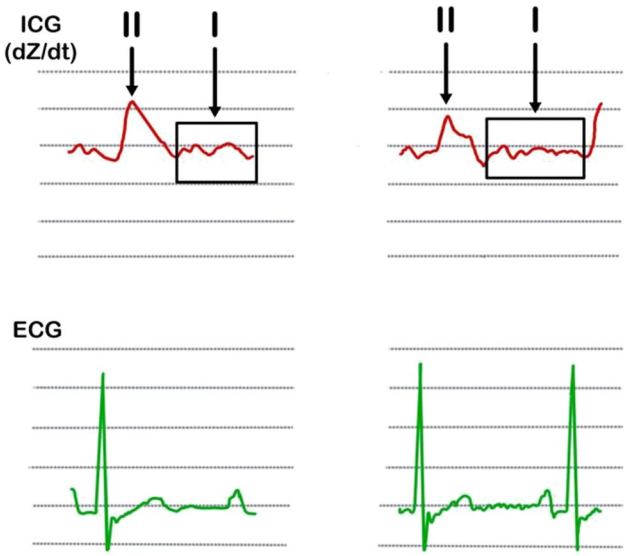
Double peak (I):



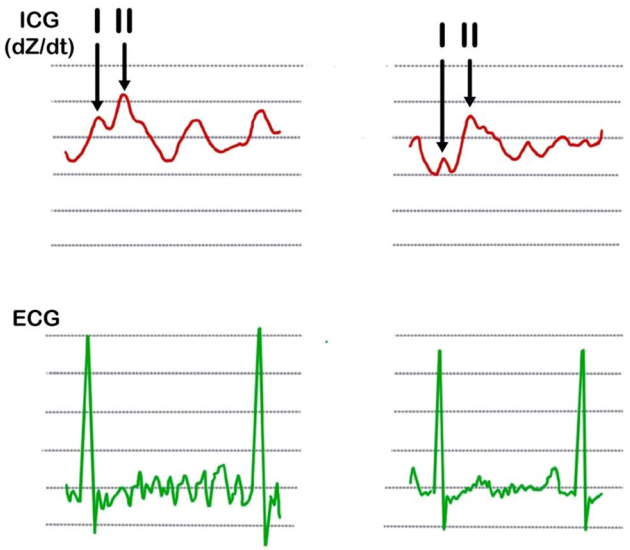
Second peak (I) which is still lower in magnitude than the first (II):



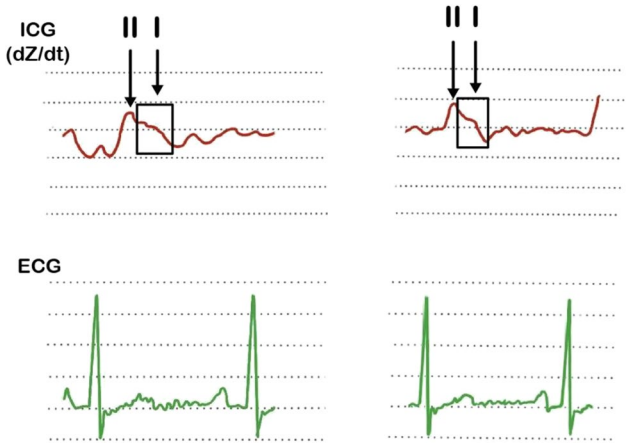
Minimal variation (I) after the peak (II):



Small notch (I) just before the peak (II):

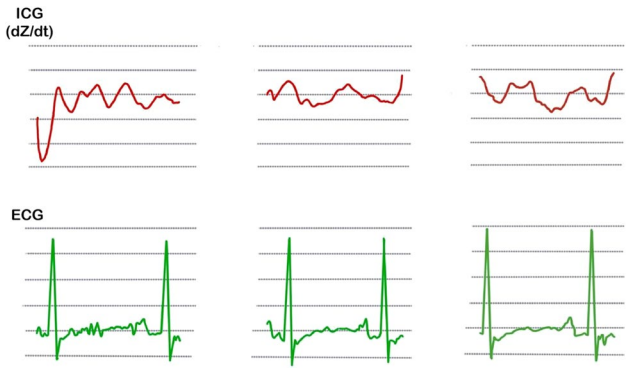


Slow decrease (I) after the peak (II):

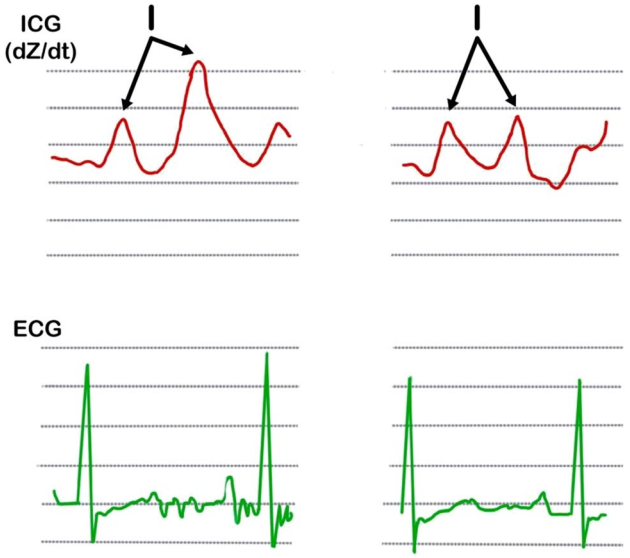


Major deviations (censoring recommended):

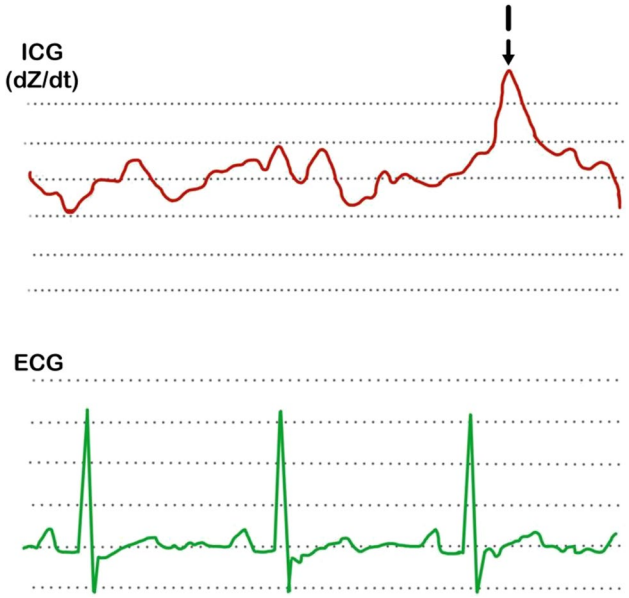
No discernible peak (or unclear which is the true peak):



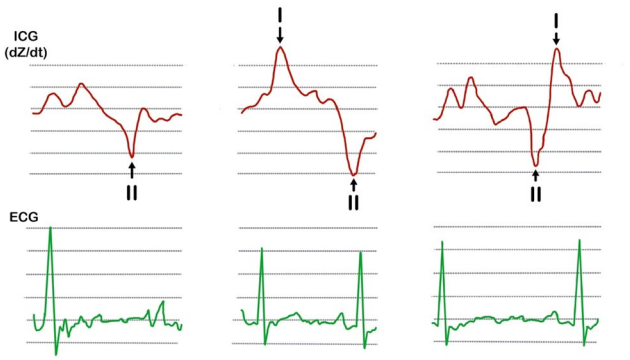
Multiple peaks of the same magnitude within one ECG cycle (I) (or greater magnitude of later peaks):



Unusually high magnitude of peak (I) in relation to other cycles:

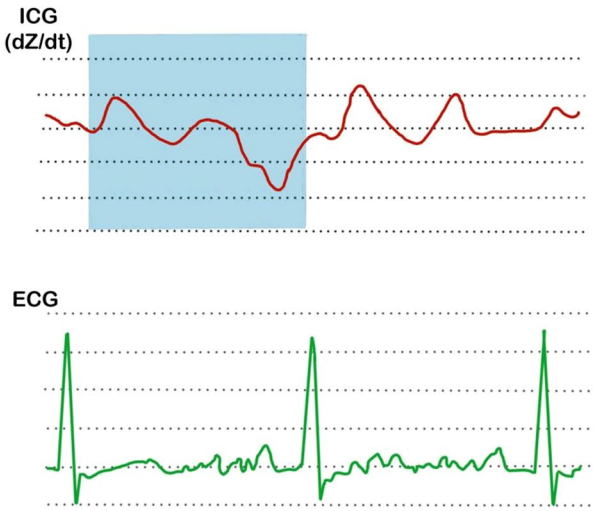


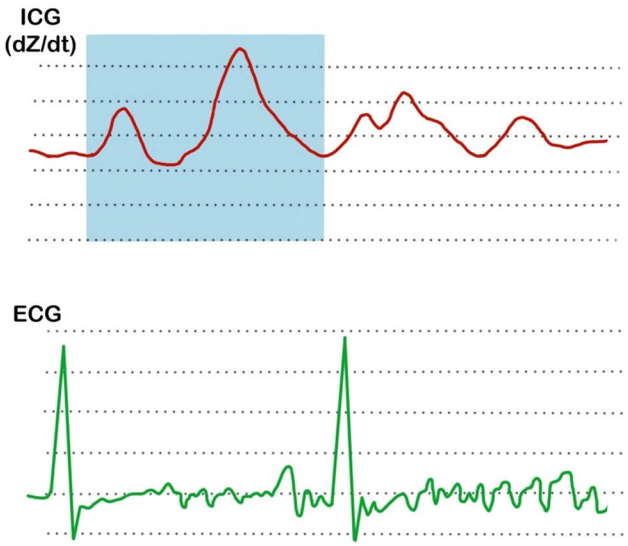
Large spikes (I) or drops (II):



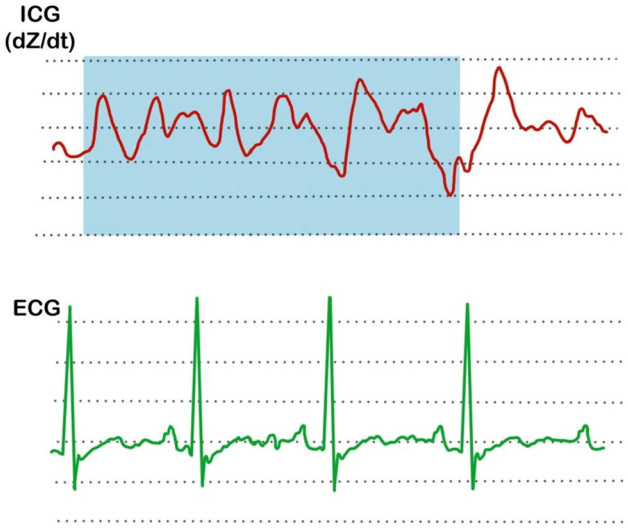
When censoring

If a cycle is disrupted or unusable, censor the entire cycle, being careful to stop just before the following ECG peak so as not to disrupt the following cycle:





Note that you will likely need to censor multiple cycles together if the disturbance is due to significant movement or some other disruption. Censor until the next clean cycle:



Appendix C: List of items to acquire

Item	Description
Computer	An up-to-date, stable computer for recording signals, pre-processing and analysis
Modular acquisition system	Modular acquisition systems integrate the incoming signals from each modular component and relay them to the computer Examples: Biopac MP series, Mindware BioNex series
ECG hardware amplifier	Amplifies the incoming electrical potential signals for digitization Examples: Biopac ECG100C amplifier, BioNex Impedance Cardiograph & GSC 2, VU-AMS
ECG electrodes	Measures electrical potential at the skin. Typically Ag/AgCl electrodes are used. Reusable and pre-gelled disposable varieties are available. With large sample sizes, the former are more cost-effective and produce less waste, but require more maintenance
ECG leadwires	Interfaces the electrodes to the hardware amplifier Refer to manufacturer's documentation to determine which cables are necessary

Item	Description
ICG hardware amplifier	Delivers current to the skin and measures voltage for determination of thoracic impedance
	Examples: Biopac NICO100C amplifier, BioNex Impedance Cardiograph & GSC 2, VU-AMS, PhysioFlow Enduro
ICG electrodes	Measures electrical potential at the skin. May be spot or band electrodes depending on your chosen configuration (see Part II—> Impedance cardiography- > ICG setup). If using spot electrodes, typically the same considerations as ECG electrodes apply (see ECG electrodes, above)
ICG leadwires	Interfaces the electrodes to the hardware amplifier
	Refer to manufacturer's documentation to determine which cables are necessary for your chosen electrode configuration (see Part II—> Impedance cardiography- > *ICG setup*)
Blood pressure measurement unit	Measures blood pressure using one of various techniques (see Part II—> Blood pressure)
	Examples: Biopac NIBP100 (A-E), CNSystems CNAP monitor, Finapres, Finometer
Transducer amplifier for blood pressure	May be necessary to interface blood pressure unit with modular research system
	Examples: Biopac DA100C; BioNex 4-Channel Transducer Amplifier
Acquisition software	For recording, processing and analysis of physiological data
	Examples: Mindware BioLab, Biopac Acqknowledge, MEAP, VU-DAMS
Miscellaneous items	
Paper towel/cleansing wipes	Assists in cleaning the electrode site to improve attachment
Electrode gel	Assists in maintaining proper connection (note: many electrodes come pre-gelled, but may require more gel)
Abrasive gel	Lightly abrades skin to remove dead skin and other debris, improving electrical contact
	Example: Nuprep Gel
Medical tape	Helps secure electrodes, particularly in the case of excess dirt, hair, or sweat
Measuring tape	Used to measure participant dimensions and distance between electrodes
Storage solutions (e.g., clips, containers, ties)	Helps maintain and store wiring and supplies to prevent damage, drying out, and save time organizing
Hard-drive backups or secure online server storage	Secures your data against hardware malfunctions, accidents etc

Acknowledgements We thank Elizabeth Page-Gould, Michael Inzlicht, Wendy Berry Mendes, Brett Q. Ford, and Jennifer E. Stellar for psychophysiological advice and support; Hause Lin for comments on the manuscript; Bentley Lo, Caroline Shi, Chiara Simeon-Spezzaferro, Daniel An, Dominic Tong, Filip Andjelkovic, Jianing Lv, Joe Hoang, Julia Won, Kristina Knox, Nicky Cheung, Octavia Yifang Weng, Rebecca Kay, and Vivian Cheng for research assistance.

Funding Preparation of this article was supported in part by the University of Toronto (University of Toronto Excellence Awards 2017 for P.R.J. and A.E.V.; Rotman School of Management Matching Research Grant for S.W.S.L.), Social Sciences and Humanities Research Council (Insight Grant 435–2017-0127 and Insight Development Grant 430–2021-00060 for S.W.S.L.), and Ontario Ministry of Research, Innovation and Science (Early Researcher Award – Round 13 for S.W.S.L.).

Data availability Not applicable. No datasets were generated or analyzed during the preparation of this tutorial.

Code availability Not applicable. No software or custom code is associated with this tutorial.

Declarations

Ethics approval Not applicable. No experiments were conducted for the preparation of this tutorial.

Consent to participate Not applicable. No experiments were conducted for the preparation of this tutorial.

Open Practices Statement Preparation of this tutorial involved no experiments; therefore no data or materials are associated with it.

Conflicts of interest/Competing interests The authors have no relevant financial or non-financial interests to disclose.

References

Ali, N., & Nater, U. M. (2020). Salivary alpha-amylase as a biomarker of stress in behavioral medicine. *International Journal of Behavioral Medicine, 27*, 337–342. https://doi.org/10.1007/s12529-019-09843-x

Árbol, J. R., Perakakis, P., Garrido, A., Mata, J. L., Fernández-Santaella, M. C., & Vila, J. (2017). Mathematical detection of aortic valve opening (B point) in impedance cardiography: A comparison of three popular algorithms. *Psychophysiology, 54*(3), 350–357. https://doi.org/10.1111/psyp.12799

Behnke, M., & Kaczmarek, L. D. (2018). Successful performance and cardiovascular markers of challenge and threat: A meta-analysis. *International Journal of Psychophysiology, 130*, 73–79. https://doi.org/10.1016/j.ijpsycho.2018.04.007

Behnke, M., Kosakowski, M., & Kaczmarek, L. D. (2020). Social challenge and threat predict performance and cardiovascular responses during competitive video gaming. *Psychology of Sport and Exercise, 46*, 101584. https://doi.org/10.1016/j.psychsport.2019.101584

Bernstein, D. P. (1986). A new stroke volume equation for thoracic electrical bioimpedance: Theory and rationale. *Critical Care Medicine, 14*(10), 904–909. https://doi.org/10.1097/00003246-198610000-00017

Berntson, G. G., Uchino, B. N., & Cacioppo, J. T. (1994a). Origins of baseline variance and the Law of Initial Values. *Psychophysiology, 31*(2), 204–210. https://doi.org/10.1111/j.1469-8986.1994.tb01042.x

Berntson, Gary G., Cacioppo, J. T., Binkley, P. F., Uchino, B. N., Quigley, K. S., & Fieldstone, A. (1994b). Autonomic cardiac control. III. Psychological stress and cardiac response in autonomic space as revealed by pharmacological blockades. Psychophysiology, 31(6), 599–608. https://doi.org/10.1111/j.1469-8986.1994b.tb02352.x

Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1993). cardiac psychophysiology and autonomic space in humans: Empirical perspectives and conceptual implications. *Psychological Bulletin, 114*(2), 296–322.

Berntson, G. G., Lozano, D. L., Chen, Y. J., & Cacioppo, J. T. (2004). Where to Q in PEP. *Psychophysiology, 41*(2), 333–337. https://doi.org/10.1111/j.1469-8986.2004.00156.x

Blascovich, J. (2008). Challenge, threat, and health. In J. Y. Shah & W. L. Gardner (Eds.), *Handbook of motivation science* (pp. 481–493). The Guilford Press.

Blascovich, J. (2013). The biopsychosocial model of challenge and threat: Reflections, theoretical ubiquity, and new directions. In *Neuroscience of prejudice and intergroup relations* (pp. 239–252). Psychology Press.

Blascovich, J., & Kelsey, R. M. (1990). Using electrodermal and cardiovascular measures of arousal in social psychological research. In C. Hendrick & M. S. Clark (Eds.), *Research methods in personality and social psychology* (pp. 45–73). Sage Publications Inc.

Blascovich, J., & Mendes, W. B. (2000). Challenge and threat appraisals: The role of affective cues. In J.P. Forgas (Ed.), *Studies in emotion and social interaction, second series. Feeling and thinking: The role of affect in social cognition* (pp. 59–82). Paris: Cambridge University Press.

Blascovich, J., & Mendes, W. B. (2010). Social psychophysiology and embodiment. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (pp. 194–227). John Wiley & Sons Inc.

Blascovich, J., Mendes, W. B., Hunter, S. B., Lickel, B., & Kowai-Bell, N. (2001). Perceiver threat in social interactions with stigmatized others. *Journal of Personality and Social Psychology, 80*(2), 253–267. https://doi.org/10.1037/0022-3514.80.2.253

Blascovich, J., Seery, M. D., Mugridge, C. A., Norris, K. R., & Weisbuch, M. (2004). Predicting athletic performance from cardiovascular indexes of challenge and threat. *Journal of Experimental Social Psychology, 40*(5), 683–688. https://doi.org/10.1016/j.jesp.2003.10.007

Blascovich, J., & Tomaka, J. (1996). The biopsychosocial model of arousal regulation. *Advances in Experimental Social Psychology, 28*, 1–51. https://doi.org/10.1016/S0065-2601(08)60235-X

Blascovich, J., Vanman, E. J., Mendes, W. B., & Dickerson, S. (2011). *Social psychophysiology for social and personality psychology*. SAGE Publications Ltd. https://doi.org/10.4135/9781446287842

Blechert, P., Peyk., Liedlgruber, M., & Wilhelm, F. H. (2016). ANSLAB: Integrated multichannel peripheral biosignal processing in psychophysiological science. *Behavior Research Methods, 48*, 1528–1545https://doi.org/10.3758/s13428-015-0665-1

Cacioppo, J. T., Berntson, G. G., Binkley, P. F., Quigley, K. S., Uchino, B. N., & Fieldstone, A. (1994). Autonomic cardiac control. II. Noninvasive indices and basal response as revealed by autonomic blockades. *Psychophysiology, 31*(6), 586–598. https://doi.org/10.1111/j.1469-8986.1994.tb02351.x

Cacioppo, J. T., & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist, 45*(1), 16–28. https://doi.org/10.1037//0003-066X.45.1.16

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. (Eds.). (2007). *The handbook of psychophysiology* (3rd ed.). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511546396

Chadha, N. J., Turner, M. J., & Slater, M. J. (2019). Investigating irrational beliefs, cognitive appraisals, challenge and threat, and affective states in golfers approaching competitive situations. *Frontiers in Psychology, 10*, 2295. https://doi.org/10.3389/fpsyg.2019.02295

Cieslak, M., Ryan, W. S., Babenko, V., Erro, H., Rathbun, Z. M., Meiring, W., Kelsey, R. M., Blascovich, J., Grafton, S. T. (2018). Quantifying rapid changes in cardiovascular state with a moving ensemble average. *Psychophysiology, 55*(4), e13018-n/a. https://doi.org/10.1111/psyp.13018

Cook, E. W., & Miller, G. A. (1992). Digital Filtering: Background and Tutorial for Psychophysiologists. *Psychophysiology, 29*(3), 350–362. https://doi.org/10.1111/j.1469-8986.1992.tb01709.x

Curtin, J. J., Lozano, D. L., & Allen, J. J. B. (2007). The psychophysiological laboratory. In J. A. Coan & J. J. B. Allen (Eds.), *Series in affective science. Handbook of emotion elicitation and assessment* (pp. 398–425). Oxford University Press.

De Veaux, R. D., Velleman, P. F., Bock, D. E., Vukov, A. M., & Wong, A. C. M. (2012). *Stats data and models (1st* (Canadian). Pearson, Canada.

Dienstbier, R. A. (1989). Arousal and physiological toughness: Implications for mental and physical health. *Psychological Review, 96*(1), 84–100.

Ditto, B., Eclache, M., & Goldman, N. (2006). Short-term autonomic and cardiovascular effects of mindfulness body scan meditation. *Annals of Behavioral Medicine, 32*(3), 227–234. https://doi.org/10.1207/s15324796abm3203_9

Folkman, S., Lazarus, R. S., Gruen, R. J., & DeLongis, A. (1986). Appraisal, coping, health status, and psychological symptoms. *Journal of Personality and Social Psychology, 50*(3), 571–579.

Frings, D., Eskisan, G., Spada, M. M., & Albery, I. P. (2015). Levels of craving influence psychological challenge and physiological reactivity. *Appetite, 84*, 161–165. https://doi.org/10.1016/j.appet.2014.10.010

Frings, D., Hurst, J., Cleveland, C., Blascovich, J., & Abrams, D. (2012). Challenge, threat, and subjective group dynamics: Reactions to normative and deviant group members. *Group Dynamics: Theory, Research, and Practice, 16*(2), 105–121. https://doi.org/10.1037/a0027504

Garrett-Ruffin, S., Hindash, A. C., Kaczkurkin, A. N., Mears, R. P., Morales, S., Paul, K., … Keil, A. (2021). Open science in psychophysiology: An overview of challenges and emerging solutions. *International Journal of Psychophysiology, 162*(January), 69–78.https://doi.org/10.1016/j.ijpsycho.2021.02.005

Gotshall, R. W., & Sexson, W. R. (1994). Comparison of band and spot electrodes for the measurement of stroke volume by the bioelectric impedance technique. *CRitical Care Medicine, 22*(3), 420–425. https://doi.org/10.1097/00003246-199403000-00011

Greene, W. A., Turetsky, B., & Kohler, C. (2000). The skeletomotor system: Surface electromyography. In J. T. Cacioppo, L. G.

Tassinary and G. Bernston (Eds.), *The handbook of psychophysiology* (2nd ed.). New York: Cambridge University Press.

Hase, A., O'Brien, J., Moore, L. J., & Freeman, P. (2019). The relationship between challenge and threat states and performance: A systematic review. *Sport, Exercise, and Performance Psychology, 8*(2), 123–144.

Helm, J. L., Miller, J. G., Kahle, S., Troxel, N. R., & Hastings, P. D. (2018). On measuring and modeling physiological synchrony in dyads. *Multivariate Behavioral Research, 53*(4), 521–543. https://doi.org/10.1080/00273171.2018.1459292

Hoyt, C. L., & Blascovich, J. (2010). The role of leadership self-efficacy and stereotype activation on cardiovascular, behavioral and self-report responses in the leadership domain. *The Leadership Quarterly, 21*(1), 89–103. https://doi.org/10.1016/j.leaqua.2009.10.007

Hurwitz, B. E., Shyu, L., Lu, C., Reddy, S. P., Schneiderman, N., & Nagel, J. H. (1993). Signal fidelity requirements for deriving impedance cardiographic measures of cardiac function over a broad heart rate range. *Biological Psychology, 36*(1–2), 3–21. https://doi.org/10.1016/0301-0511(93)90076-K

Jamieson, J. P., Nock, M. K., & Mendes, W. B. (2012). Mind over matter: Reappraising arousal improves cardiovascular and cognitive responses to stress. *Journal of Experimental Psychology: General, 141*(3), 417–422. https://doi.org/10.1037/a0025719

Jänig, W. (2006). *The Integrative Action of the Autonomic Nervous System: Neurobiology of Homeostasis*. Cambridge University Press.

Jennings, J. R., & Gianaros, P. J. (2007). Methodology. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (3rd ed., pp. 812–833). Cambridge University Press. https://doi.org/10.1017/CBO9780511546396.034

Juster, R. P., McEwen, B. S., & Lupien, S. J. (2010). Allostatic load biomarkers of chronic stress and impact on health and cognition. *Neuroscience & Biobehavioral Reviews, 35*(1), 2–16. https://doi.org/10.1016/j.neubiorev.2009.10.002

Kassam, K. S., Koslov, K., & Mendes, W. B. (2009). Decisions under distress: Stress profiles influence anchoring and adjustment. *Psychological Science, 20*(11), 1394–1399. https://doi.org/10.1111/j.1467-9280.2009.02455.x

Kelsey, R. M., Blascovich, J., Tomaka, J., Leitten, C. L., Schneider, T. R., & Wiens, S. (1999). Cardiovascular reactivity and adaptation to recurrent psychological stress: Effects of prior task exposure. *Psychophysiology, 36*(6), 818–831. https://doi.org/10.1111/1469-8986.3660818

Kelsey, R. M., & Guethlein, W. (1990). An evaluation of the ensemble averaged impedance cardiogram. *Psychophysiology, 27*(1), 24–33. https://doi.org/10.1111/j.1469-8986.1990.tb02173.x

Kelsey, R. M., Soderlund, K., & Arthur, C. M. (2004). Cardiovascular reactivity and adaptation to recurrent psychological stress: Replication and extension. *Psychophysiology, 41*(6), 924–934. https://doi.org/10.1111/j.1469-8986.2004.00245.x

Kirschbaum, C., & Hellhammer, D. (1989). Salivary Cortisol in Psychobiological Research : An Overview. *Biological Psychology/Psychopharmacology*, *22*, 150–169. https://doi.org/10.1159/000118611

Kligfield, P., Gettes, L. S., Bailey, J. J., Childers, R., Deal, B. J., Hancock, E. W., … Wagner, G. S. (2007). Recommendations for the standardization and interpretation of the electrocardiogram: Part I: The electrocardiogram and its technology: A scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society. *Circulation*, *115*(10), 1306–1324.https://doi.org/10.1161/CIRCULATIONAHA.106.180200

Kubicek, W. G., Patterson, R. P., & Witsoe, D. A. (1970). Impedance cardiography as a noninvasive method of monitoring cardiac function and other parameters of the cardiovascular system. *Annals of the New York Academy of Sciences, 170*(2), 724–732. https://doi.org/10.1111/j.1749-6632.1970.tb17735.x

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer publishing company.

Lenis, G., Pilia, N., Loewe, A., Schulze, W. H. W., & Dössel, O. (2017). Comparison of Baseline Wander Removal Techniques Considering the Preservation of ST Changes in the Ischemic ECG: A Simulation Study. *Computational and Mathematical Methods in Medicine*, *2017*.https://doi.org/10.1155/2017/9295029

Lindsey, M. L., Gray, G. A., Wood, S. K., & Curran-Everett, D. (2018). Statistical considerations in reporting cardiovascular research. *American Journal of Physiology-Heart and Circulatory Physiology, 315*(2), H303–H313. https://doi.org/10.1152/ajpheart.00309.2018

Llabre, M. M., Spitzer, S. B., Saab, P. G., Ironson, G. H., & Schneiderman, N. (1991). The reliability and specificity of delta versus residualized change as measures of cardiovascular reactivity to behavioral challenges. *Psychophysiology, 28*(6), 701–711. https://doi.org/10.1111/j.1469-8986.1991.tb01017.x

Luo, S., & Johnston, P. (2010). A review of electrocardiogram filtering. *Journal of Electrocardiology, 43*(6), 486–496. https://doi.org/10.1016/j.jelectrocard.2010.07.007

Maier, K. J., Waldstein, S. R., & Synowski, S. J. (2003). Relation of cognitive appraisal to cardiovascular reactivity, affect, and task engagement. *Annals of Behavioral Medicine, 26*(1), 32–41. https://doi.org/10.1207/S15324796ABM2601_05

McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine, 338*(3), 171–179. https://doi.org/10.1056/NEJM199801153380307

McEwen, B. S. (2004). Protection and damage from acute and chronic stress: Allostasis and allostatic overload and relevance to the pathophysiology of psychiatric disorders. *Annals of the New York Academy of Sciences, 1032*(1), 1–7. https://doi.org/10.1196/annals.1314.001

Meidert, A. S., & Saugel, B. (2018). Techniques for non-invasive monitoring of arterial blood pressure. *Frontiers in Medicine, 4*, 1–6. https://doi.org/10.3389/fmed.2017.00231

Mendes, W. B. (2009). Assessing autonomic nervous system activity. In E. Harmon-Jones & J. S. Beer (Eds.), *Methods in social neuroscience* (pp. 118–294). Guilford Press.

Mendes, W. B., Blascovich, J., Hunter, S. B., Lickel, B., & Jost, J. T. (2007a). Threatened by the unexpected: Physiological responses during social interactions with expectancy-violating partners. *Journal of Personality and Social Psychology, 92*(4), 698–716. https://doi.org/10.1037/0022-3514.92.4.698

Mendes, W., Blascovich, J., Lickel, B., & Hunter, S. (2002). Challenge and threat during social interactions with white and black men. *Personality and Social Psychology Bulletin, 28*(7), 939–952. https://doi.org/10.1177/014616720202800707

Mendes, W. B., Blascovich, J., Major, B., & Seery, M. (2001). Challenge and threat responses during downward and upward social comparisons. *European Journal of Social Psychology, 31*(5), 477–497. https://doi.org/10.1002/ejsp.80

Mendes, W., Gray, H. M., Mendoza-Denton, R., Major, B., & Epel, E. S. (2007b). Why egalitarianism might be good for your health: Physiological thriving during stressful intergroup encounters. *Psychological Science, 18*(11), 991–998. https://doi.org/10.1111/j.1467-9280.2007.02014.x

Mohapatra, S. N. (1981). *Non-invasive cardiovascular monitoring by electrical impedance technique*. Pitman Medical.

Morgan, E. (2017, December 6). *Improving Data Quality: General Guidelines*. MindWare Technologies LTD. Support Portal. https://support.mindwaretech.com/2017/12/improving-data-quality-general-guidelines/

Obrist, P. (1981). *Cardiovascular psychophysiology: A perspective*. Plenum.

Patil, P., Peng, R. D., & Leek, J. T. (2016). What should researchers expect when they replicate studies? A statistical view of replicability in psychological science. *Perspectives on Psychological Science, 11*(4), 539–544. https://doi.org/10.1177/1745691616646366

Peters, B. J., & Jamieson, J. P. (2016). The consequences of suppressing affective displays in romantic relationships: A challenge and threat perspective. *Emotion, 16*(7), 1050–1066. https://doi.org/10.1037/emo0000202

Peters, B. J., Overall, N. C., & Jamieson, J. P. (2014). Physiological and cognitive consequences of suppressing and expressing emotion in dyadic interactions. *International Journal of Psychophysiology, 94*(1), 100–107. https://doi.org/10.1016/j.ijpsycho.2014.07.015

Qu, M., Zhang, Y., Webster, J. G., & Tompkins, W. J. (1986). Motion artifact from spot and band electrodes during impedance cardiography. *IEEE Transactions on Biomedical Engineering, BME-33*(11), 1029–1036. https://doi.org/10.1109/TBME.1986.325869

Quigley, K. S., Barrett, L. F., & Weinstein, S. (2002). Cardiovascular patterns associated with threat and challenge appraisals: A within-subjects analysis. *Psychophysiology, 39*(3), 292–302. https://doi.org/10.1111/psyp.13454

Reifman, A. & Keyton, K. (2010). Winsorize. In N. J. Salkind (Ed.), *Encyclopedia of research design* (pp. 1637–1638). Thousand Oaks, CA: SAGE Publications, Inc. https://doi.org/10.4135/9781412961288.n502

Ruppert, D. (2006). Trimming and winsorization. In S. Kotz, C. B. Read, N. Balakrishnan, B. Vidakovic and N. L. Johnson (Eds.), *Encyclopedia of statistical sciences.* https://doi.org/10.1002/0471667196.ess2768.pub2

Ryan, W., Conigrave, J. H., Basarkod, G., Ciarrochi, J., & Sahdra, B. K. (2019). When is it good to use wristband devices to measure HRV?: Introducing a new method for evaluating the quality of data from photophlethysmography-based HRV devices. https://doi.org/10.31234/osf.io/t3gdz

Seery, M. D. (2011). Challenge or threat? Cardiovascular indexes of resilience and vulnerability to potential stress in humans. *Neuroscience & Biobehavioral Reviews, 35*(7), 1603–1610. https://doi.org/10.1016/j.neubiorev.2011.03.003

Seery, M. D. (2013). The biopsychosocial model of challenge and threat: Using the heart to measure the mind. *Social and Personality Psychology Compass, 7*(9), 637–653. https://doi.org/10.1111/spc3.12052

Seery, M. D., Weisbuch, M., & Blascovich, J. (2009). Something to gain, something to lose: The cardiovascular consequences of outcome framing. *International Journal of Psychophysiology, 73*(3), 308–312. https://doi.org/10.1016/j.ijpsycho.2009.05.006

Seery, M. D., Weisbuch, M., Hetenyi, M. A., & Blascovich, J. (2010). Cardiovascular measures independently predict performance in a university course. *Psychophysiology, 47*(3), 535–539. https://doi.org/10.1111/j.1469-8986.2009.00945.x

Semmlow, J. (2018). The Big Picture: Bioengineering Signals and Systems. In J. Semmlow (Ed.), *Biomedical Engineering, Circuits, Signals and Systems for Bioengineers (Third Edition)* (pp. 3–50). Academic Press. https://doi.org/10.1016/B978-0-12-809395-5.00001-1

Sherwood, A., Allen, M. T., Fahrenberg, J., Kelsey, R. M., Lovallo, W. R., & van Doornen, L. J. P. (1990). Methodological guidelines for impedance cardiography. *Psychophysiology, 27*(1), 1–23. https://doi.org/10.1111/j.1469-8986.1990.tb02171.x

Sherwood, A., Royal, S. A., Hutcheson, J. S., & Turner, J. R. (1992). Comparison of impedance cardiographic measurements using band and spot electrodes. *Psychophysiology, 29*(6), 734–741. https://doi.org/10.1111/j.1469-8986.1992.tb02051.x

Sillars, A. A., & Davis, E. L. (2018). Children's challenge and threat appraisals vary by discrete emotion, age, and gender. *International Journal of Behavioral Development, 42*(5), 506–511. https://doi.org/10.1177/0165025417739178

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Stanley, T. D., Carter, E. C., & Doucouliagos, H. (2018). What meta-analyses reveal about the replicability of psychological research. *Psychological Bulletin, 144*(12), 1325–1346. https://doi.org/10.1037/bul0000169

Stern, R. M., Ray, W. J., & Quigley, K. S. (2001). Some Basic Principles of Psychophysiology. In Stern, R. M., Ray, W. J., & Quigley, K. S. (Eds.), Psychophysiological recording. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195113594.003.0005

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Allyn & Bacon/Pearson Education.

Thorson, K. R., West, T. V., & Mendes, W. B. (2018). Measuring physiological influence in dyads: A guide to designing, implementing, and analyzing dyadic physiological studies. *Psychological Methods, 23*(4), 595–616. https://doi.org/10.1037/met0000166

Tomaka, J., Blascovich, J., Kelsey, R. M., & Leitten, C. L. (1993). Subjective, physiological, and behavioral effects of threat and challenge appraisal. *Journal of Personality and Social Psychology, 65*(2), 248–260. https://doi.org/10.1037/0022-3514.65.2.248

Tomaka, J., Blascovich, J., Kibler, J., & Ernst, J. M. (1997). Cognitive and physiological antecedents of threat and challenge appraisal. *Journal of Personality and Social Psychology, 73*(1), 63–72. https://doi.org/10.1037//0022-3514.73.1.63

Tomaka, J., & Magoc, D. (2021). Personality antecedents of challenge and threat appraisal. *Stress and Health*, 1–10.https://doi.org/10.1002/smi.3028

Vick, S. B., Seery, M. D., Blascovich, J., & Weisbuch, M. (2008). The effect of gender stereotype activation on challenge and threat motivational states. *Journal of Experimental Social Psychology, 44*(3), 624–630. https://doi.org/10.1016/j.jesp.2007.02.007

VU-AMS Ambulatory Monitoring System. (n.d.). *Impedance scoring.* http://www.vu-ams.nl/support/previous-ams-versions/manuals/amsimp/impedance-scoring/

Weisbuch-Remington, M., Mendes, W. B., Seery, M. D., & Blascovich, J. (2005). The nonconscious influence of religious symbols in motivated performance situations. *Personality and Social Psychology Bulletin, 31*(9), 1203–1216. https://doi.org/10.1177/0146167205274448

Widmann, A., Schröger, E., & Maess, B. (2015). Digital filter design for electrophysiological data – a practical approach. *Journal of Neuroscience Methods, 250*, 34–46. https://doi.org/10.1016/j.jneumeth.2014.08.002

Wormwood, J. B., Khan, Z., Siegel, E., Lynn, S. K., Dy, J., Barrett, L. F., & Quigley, K. S. (2019). Physiological indices of challenge and threat: A data-driven investigation of autonomic nervous system reactivity during an active coping stressor task. *Psychophysiology, 56*(12), 1–18. https://doi.org/10.1111/psyp.13454

Wright, R. A., & Kirby, L. D. (2001). Effort determination of cardiovascular response: An integrative analysis with applications in social psychology. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 33, pp. 255–307). Academic Press.