The version of record is available at https://doi.org/10.1108/CI-02-2022-0040. The full citation is as follows: Newton, S. (2022), "Measuring the perceptual, physiological and environmental factors that impact stress in the construction industry", Construction Innovation, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/CI-02-2022-0040

Construction Innovation: Information, Process, Management

Measuring the perceptual, physiological and environmental factors that impact stress in the construction industry

Journal:	*Construction Innovation: Information, Process, Management*
Manuscript ID	CI-02-2022-0040.R1
Manuscript Type:	Research Article
Keywords:	Biometric, Electrodermal Activity, Psychometric, Pulse Rate Variability, Stress, Anxiety

Measuring the perceptual, physiological and environmental factors that impact stress in the construction industry.

Abstract

Purpose: The aim of this study is to highlight and demonstrate how the study of stress and related responses in construction can best be measured and benchmarked effectively.

Design/methodology/approach: A range of perceptual and physiological measures are obtained across different time periods and during different activities in a fieldwork setting. Differences in the empirical results are analysed and implications for future studies of stress discussed.

Findings: Results strongly support the use of multiple psychometrics and biosensors whenever biometrics are included the study of stress. Perceptual, physiological and environmental factors are shown to all act in concert to impact stress. Strong conclusions on the potential drivers of stress should then only be considered when consistent results apply across multiple metrics, time periods and activities.

Originality: First study to focus explicitly on demonstrating the need for multiple research instruments and settings when studying stress or related conditions in construction.

Research limitations/implications: Stress is an incredibly complex condition. This study demonstrates why many current applications of biosensors to study stress in construction are not up to the task, and provides empirical evidence on how future studies can be significantly improved.

1 Introduction

1.1 Study objectives

The construction industry has long been associated with producing a stressful workplace – high-pressure project schedules, hazardous work environments, unhelpful social stigmas, and complex work practices (Campbell, 2006; Chan et al., 2020). The stress of construction work is not confined to the construction site. Stress is recognised as an issue across the construction industry, including for construction professionals, regulators, trainers and many others associated with construction work (Kamardeen, 2022). In various contexts, high levels of occupational stress have been shown to result in lower productivity, poor physical wellbeing, and a significant increase in work-related mental illnesses (Langdon and Sawang, 2018; Hagermoser Sanetti et al., 2020).

Traditionally, the primary means used to study stress in construction has relied on self-reported psychometric tests of perceived stress levels. Psychometric stress tests have certainly provided important insight into the causes, management and mitigation of stress in various aspects of construction (see for example, Love et al., 2010). Nevertheless, sole reliance on the subjective recall of self-reporting individual perceptions can introduce significant issues of misperception and the limitations of perceptual bias (Textor, 2019). Furthermore, as the aetiology of stress is better understood, a complex of factors, levels, characteristics and modes has been associated with stress – a complex that current psychometric testing alone is often not able to differentiate (Crosswell and Lockwood, 2020).

Recent developments in the field of neuroscience and biometric sensor technologies have promoted the application of affordable biosensors to study the physiology of the stress response more directly and objectively (Jebelli et al., 2019). These biosensors are being applied increasingly to study stress responses in general (Sampson and Koh, 2020). There is also now a rapidly expanding literature that employs biosensors to study stress and related emotional and cognitive responses to risk, workload, fatigue, awareness, etc., in a variety of construction settings (Saedi et al., 2022). That trend is set to accelerate as biosensors become more generally accessible, and biometric data processing tools are improved. However, biosensors have their own attendant problems for the measurement of biomarkers in general, and for stress in particular (Samson and Koh, 2020). Stress is just one potential in a complex of physiological responses that can all register in very similar ways using common biosensor monitors – making it difficult to distinguish stress from various other factors, or to distinguish those other factors from each other (Geršak, 2020).

Significantly, stress responses are often (incorrectly) conflated with anxiety responses. Where stress is considered as an external trigger of emotional response, anxiety is a more persistent internal driver of emotion. Indeed, stress and anxiety are complementary aspects of the same stress/anxiety complex. This makes it extremely problematic to consider either in isolation from the other when measuring relevant biometrics (Bystritsky and Kronemyer, 2014). Furthermore, an expansive range of personal and environmental factors are known to mediate how a particular individual, in a particular situation, manifests their biomarkers (Crosswell and Lockwood, 2020).

This study is motivated by the number of otherwise sound recent studies in construction that employ biosensors to study stress and related emotional responses, but fall short of incorporating the necessary combination of perceptual, physiological and environmental measures. It is all three of these factors, acting in concert, that influence stress. All three factors are therefore implicated in the measurement of stress and the various other emotional and cognitive responses increasingly being investigated with biosensor technologies (Jebelli et al., 2019; Chae et al., 2021). Nevertheless, there are construction studies that omit to sample (and thereby to account for) the emotional disposition of individual participants. For example, studies that do not undertake relevant psychometric testing of the participants (Hwang et al., 2016; Jeon and Cai, 2021; Nwaogu and Chan, 2021). There are construction studies that omit to measure the range of appropriate physiological biomarkers. For example, because they focus on a single metric (Choi et al., 2019; Chen and Tserng, 2022), measure at an insufficient sample rate (Jebelli et al., 2018), or for an insufficient period of time (Ke et al., 2021). There are construction studies that omit to provide an ecologically valid setting for the research. For example, because they rely on laboratory-based tests of site conditions (Jeon and Cai, 2021; Ke et al., 2021), or because they employ contrived or otherwise poorly simulated construction site and office settings (Jeon and Cai, 2021; Chae et al., 2021).

The purpose of this study is to demonstrate that both psychometric and biometric instruments are required, collectively, and applied in ecologically valid settings, to make adequate sense of the stress responses of individuals (and by implication other related emotional and cognitive responses). Further, that whenever multiple research instruments are applied to the same study, results are generally more nuanced. Particular care is then needed to ensure that conclusions are only drawn when consistent results across instruments are found.

The focus of the study is therefore on the extent to which the outcomes of a stress study can vary depending on the emotional disposition of the participants, the choice of biomarkers being measured, and environmental changes to the temporal and activity settings. That variation is demonstrated using multiple psychometric and biometric measures across different time periods with different activities, compared between two participants in an identical fieldwork setting. The data is derived from a study of two construction management educators in a classroom (fieldwork) setting, team teaching across 3 days. Whilst each construction setting may be different, the primary consideration for this study is to show that quality data can be captured outside of the laboratory, in an ecologically valid (fieldwork) setting. The study demonstration then measures, analyses and compares a range of selected emotional and physiological biomarkers from two participants in the exact same situation, across time and in a variety of environmental settings. Discrepancies in the results highlight the potential variability between each metric and setting.

Previous studies of stress in the construction industry have sought to include a combination of perceptual, physiological and environmental factors (see, for example, Lee et al., 2017). However, this study is the first to incorporate such an extensive range of psychometric and biometric instruments, across time, and with markedly different activities. It is the first to focus explicitly on highlighting the need for multiple research instruments and settings when studying stress or other related emotional and cognitive responses in construction. This empirical demonstration of how perception, physiology and environmental factors act in concert, and should therefore all be represented and accounted for in the study of stress, is of particular significance given the accelerating application of biosensors to construction settings.

1.2 Stress and anxiety

In general, stress is most often used as a term to represent the mental or emotional strain caused when an individual is unable to respond effectively to the challenges or adverse conditions being encountered (Fink, 2017). Stressful challenges or "stressors" are the discernible and discrete events or circumstances that can trigger a stress response. Stressors can vary by type (from physical to social to cognitive) and significance (from the mundane to the catastrophic). Different individuals may respond differently to the same stressor depending on how it presents (the duration, circumstance, how it is perceived, etc.). Significantly, stressors can manifest in a variety of forms, from acute or otherwise traumatic episodes of limited duration, as a single point individual life event, or as a chronic and more persistent disruption to the physiological steady-state (or 'homeostasis') over time (Crosswell and Lockwood, 2020). In this way, even otherwise relatively mundane, so-called "daily hassles" can accumulate over time, to overwhelm the capacity of an individual to respond effectively. A stressor is therefore anything that causes a biological or behavioural correction to the threat of a homeostasis imbalance (such anticipatory adaptation is called the 'allostatic load') (Samson and Koh, 2020).

The stress response of an individual depends on the physiological, cognitive, behavioural and emotional resources and options available to them before, during and after the exposure to stressors (Cohen et al., 2016). That dependency accounts for the manifest potential for stress to cause both serious physical and mental health issues, as the coping mechanisms themselves (the heightened allostatic load) can damage the very biological systems on which they depend (Epel et al., 2018). The stress responses also range across multiple levels (social, psychological, physiological, etc.) and involve several related but otherwise discrete biological systems. Collectively, this not only means that stress responses are complex and multifarious, but also renders them as rather amorphous and difficult to 'locate' in specific physiological systems. The diversity also demands a range of research instruments and multiple metrics in a variety of settings, are necessary to enable meaningful conclusions to be drawn from studies of stress, stress-related behavioural reactions, or stress-related health outcomes (Crosswell and Lockwood, 2020).

Like stress, anxiety typically arises in response to a perceived environmental threat. Anxiety however comes primarily from the mental response to a fear and/or anticipation of the potential threat of an imminent or uncertain outcome (Julian, 2011). Anxiety, whilst prompted by environmental factors and manifested as allostatic load, is entirely more psychological in nature than is the case for stress. Nevertheless, the anxiety responses overlap and intertwine considerably with regular stress responses. The two sets of responses share the vast majority of significant response features and biomarkers, and indeed there is a reciprocal and dynamic relationship between stress and anxiety (Epel et al., 2018). Stress can lead to anxiety. Anxiety will condition stress. This cumulative and compounding relationship between stress and anxiety renders the independent measurement of stress even more intractable. It requires the factors associated with anxiety, and the problematic cognitive, affective and behavioural processing that accompanies anxiety, to be factored in to any consideration of stress (Bystrisky and Kronemyer, 2020). Further, the inseparable stress/anxiety complex is of particular significance to emerging biometric measurement, because similar neurobiological mechanisms (and thereby the same biomarkers) are implicated in both stress and anxiety responses (Geršak, 2020).

2 Identifying meaningful measurement instruments

2.1 Identifying relevant psychometrics

According to the prevailing transactional theories of emotions (Lazarus and Folkman 1987), the experience of stress or anxiety involves an actual, physiological arousal (precipitated by stressors) in necessary combination with a cognitive appraisal (testing beliefs about the nature and cause of the arousal). To experience stress or anxiety a person must first perceive a situation or circumstance to ‘be’ stressful or anxious (Bystritsky and Kronemyer, 2020). Thus, the measurement of perceptions is essential to the study of stress. The measurement of perceptions generally requires the construction and validation of assessment instruments such as interviews and questionnaires.

There is currently no universally recognised standard for stress or anxiety evaluation (Kim et al., 2018). However, multiple and various psychometric instruments have been developed and tested which have helped better clarify the important study dimensions and parameters of stress and anxiety (Fink, 2017). For example, both stress and anxiety (at least in general) are best considered to have two broad dimensions: state/phasic conditions, which reflect the presence and severity of current symptoms; and trait/tonic conditions, which reflect the generalised propensity of an individual towards stress and/or anxiety. Further, both stress and anxiety are heavily influenced by a variety of individual and situational factors (such as age, health, temperature, activity, and environmental distractions). There are also various forms of stress (acute, chronic, post-traumatic, etc.) and anxiety (social, specific phobias, panic attacks, obsessive compulsive disorders, etc), including generalised forms of both (Karatsoreos, 2018).

There is now an extensive array of established psychometric survey instruments developed specifically to measure perceived stress and anxiety in various contexts and with various demographics. Arguably the most widely used stress survey instrument is the Perceived Stress Scale (PSS), which measures generalised stress (Ribeiro Santiago et al., 2020). The typical PSS comprises 14 statements, each with a five-point response scale. The statements evaluate a two-factor structure of perceived stress (negative) and perceived coping (positive). Participants are asked about their feelings and thoughts during a specified period (the past 24 hours, week, month or year, depending on the study purpose). In each question, the participant is asked how often they have felt or thought a certain way. The PSS score (the higher the value the higher the perceived level of stress) is calculated by summing the stress scores and subtracting the coping scores. Whilst the PSS remains the most widely used instrument, a recent validation study concluded that there were critical concerns with the validity of several questions; discouraged the combination of the stress and control sub-scores into a single PSS score; and identified measurement bias when scores for each question are weighted equally (Ribeiro Santiago et al., 2020). The findings would indicate that PSS requires further validation in other large sample general populations particular to other countries and cultures. However, at this stage the PSS remains the most robust measure of perceived general stress (Lee, 2012).

Anxiety also requires a raft of tailored questionnaire instruments to measure the many variations in diagnostic criteria and specific anxiety disorders (Julian, 2011). However, for generalised anxiety measurement the State–Trait Inventory of Cognitive and Somatic Anxiety (STICSA) is a commonly administered self-report instrument of increasing popularity (Grös et al., 2007). The STICSA consists of two identical, 21-item scales: a state scale that measures mood at that moment in time; and a trait scale that measures mood in general. The, 21 items comprise two subscales: 10 items particular to cognition and how participants think about anxiety; and 11 items particular to somatic anxiety and how people feel about anxiety. Items are measured on a 4-point scale ranging from 1 (not at all) to 4 (very much so). A recent validity study of STICSA (Styck et al., 2020) highlighted potential convergence problems, but concluded that these correlations are inescapable in a multidimensional psychological construct. The use of STICSA may best be justified when applied to simply structured (more homogenous) populations (Styck et al., 2020).

Psychometrics offer important insight into the disposition of a study participant to stress and anxiety, because they sample the cognitive perceptions that test and determine how such arousals are processed. There are well-established psychometric instruments for measuring general stress and general anxiety in small populations. However, further development and extension of alternative instruments may be required when considering stress in specific, complex and/or large sample populations, or when more specific forms of anxiety are being studied.

2.2 Identifying relevant biometrics

Stress and anxiety responses are promoted through nearly every system of the body. However, two systems are most frequently used for measurement: the autonomic nervous system (ANS) and the hypothalamic-pituitary-adrenal (HPA) axis. The ANS is an involuntary and rapidly operating (sub-second timeframes) mechanism. It primarily regulates through two very distinct, often antagonistic but ultimately complementary pathways: the sympathetic nervous system (SNS), and the parasympathetic nervous system (PNS). The SNS actively promotes immediate change (increase) in response to a perceived threat or fear. The PNS actively seeks to calm responses and maintain resting functions over the longer term. The PNS can, however, respond immediately to supress itself and thus enable the impact of the SNS to be less constrained and more impactful (Krebig and Gendolla, 2014).

The ANS activity is most often measured through associated changes in the heart rate (HR), heart rate variability (HRV), and pulse rate variability (PRV). Various biometric technologies are able to provide HR, HRV and/or PRV readings. PhotoPlethysmoGraphy (PPG) is often favoured because it provides a simple, single measure of the blood volume pulse (BVP) from which an accurate HR and PRV can be determined efficiently. Direct measurement of the HRV requires an Electrocardiogram (ECG) monitor, to record the electrical signals of the heart. HRV is then considered the better measure of cardiac rhythm, and is often favoured in clinical studies of workplace stress (Järvelin-Pasanen et al., 2018). Whilst the issue lacks consensus (Mejía-Mejía et al., 2020a), PRV is often used as a surrogate for HRV. HRV/PRV is a common choice of biomarker in construction studies that employ biometrics (Lee et al., 2017; Nwaogu and Chan, 2021), and is generally preferred for the assessment of ANS responses more broadly (Mejía-Mejía et al., 2020b). In any event, HRV/PRV is ultimately more of a relative measure of an individual's capacity to respond effectively to stress and other ANS stimulations over time. It typically requires an individual baseline measure to be determined over a period of days, and even then, fluctuates constantly depending on multiple environmental factors and activity levels (Chalmers et al., 2022).

ANS activity is also increasingly measured by monitoring changes in ElectroDermal Activity (EDA). EDA provides a measure of the electrical resistance/conductance associated with skin sweat secretion. One involuntary consequence of the ANS being stimulated, is the tiny volumes of skin sweat secretion measured by EDA sensors. EDA is an increasingly popular measure of ANS responses because it is immediate, non-intrusive and can be monitored continuously using low-cost and stable devices. As with other biomarkers of ANS stimulation however, EDA requires a baseline measure for each individual to be established over several hours.

The HPA axis relies on a slower acting system of various chemical messengers (hormones) that travel through the bloodstream and induce specific functional responses from specific organs and tissues in the body. There are several important hormones involved in stress responses, but arguably the best direct indicator of HPA activity relevant to stress is Cortisol (Ali and Nater, 2020). This is because Cortisol is directly implicated in stress management, and Cortisol levels can be measured non-invasively in several readily-available bodily fluids (including sweat and saliva) (Samson and Koh, 2020).

Ultimately, managing both the ANS and HPA is the brain. Brain activity can also then offer a direct measurement of stress and anxiety responses. Brain activity during cognitive tasks is most typically monitored using fMRI (functional Magnetic Resonance Imaging) technologies. fMRI measures the changes in blood oxygenation and blood flows through the brain that occur in response to neural activity. However, fMRI has prohibitive limitations for the study of stress and anxiety specifically, due to the size and stress-inducing nature of the fMRI technologies themselves. fMRI testing is both loud and claustrophobic. Recent advances in non-invasive fNIRS (functional Near-InfraRed Spectroscopy) measure secondary blood oxygenation and flow at the cortical surface. fNIRS offers a much-improved option over fMRI in terms of mobility and temporal resolution, making fNIRS the preferred technology for stress and anxiety measurement in ecologically valid (ie. fieldwork) settings (Quaresima and Ferrari, 2019).

Prior to fNIRS the most viable, and still the most popular, technology for mobile brain activity imaging is ElectroEncephaloGraphy (EEG). EEG measures the electrical activity of the brain directly, providing significantly more temporal resolution (milliseconds) than the fNIRS, but having the distinct disadvantage of introducing a substantial extent of signal noise artefacts. Significant signal noise artefacts can be introduced through nothing more than eye movement and other facial muscle activities, cardiac activity through blood pulses, and even the extrinsic environmental electrical activity of mobile phones, computers and other nearby electronic devices (Jiang et al., 2019). Signal noise artefacts are especially problematic when the EEG device only records a small number of channels – less than, say, 64 electrodes. It is noteworthy that the consumer-grade, wearable EEG devices typically being used by academic researchers currently (see for example, Jeon and Cai, 2021; Ke et al., 2021) are limited to just 14 channels/electrodes. A recent assessment by Wexler and Thibault (2019) concluded that such consumer-grade EEG devices do not record brain activity with sufficient validly or reliably to accurately reflect the mental states they claim to measure.

Biosensors are of growing interest to the study of stress and anxiety in construction because they offer a more objective measure of the relevant biomarkers. This means biosensors offer a critical complement to psychometric measurement. It does not mean that biosensors offer a valid alternative or stand-alone measurement for stress. In any event, there is no single biomarker for stress or anxiety. Multiple forms of biosensors, measuring multiple biometrics, are necessary to make sense of physiological responses to stress.

When considering the ANS response, the most promising biometric measures are provided by PPG and EDA biosensors. Alternatively, for HPA activity, Cortisol levels also provide an effective indicator. When considering brain activity more directly, fNIRS technology offers future promise. However, the use of many consumer-grade EEG devices is to be discouraged – at least until a significantly higher density of signals (minimum 64 channels in general, and increasing to 125 channels when task specific placement of electrodes is not possible, Lau et al., 2012) and/or greatly improved signal quality is achieved.

3 Study Method

3.1. Context and procedures

The aim of this study is to highlight and demonstrate how the study of stress and anxiety in construction can best be measured and benchmarked effectively. A demonstration study is presented to highlight the key characteristics for an effective study of stress: perceptual and physiology research instruments should be used in combination, and not independently; multiple psychometric and biometric measurements are necessary to make sense of the stress/anxiety complex; data collection must include a sufficient richness of data points, recorded over a sufficient timeframe, at sufficient density, and across a sufficient range of participant activities to warrant strong conclusions; and the stress study environment, especially for construction-related studies, heavily prioritises the ecological validity of relevant fieldwork settings over laboratory-based settings.

A range of perceptual and physiological research instruments are considered and incorporated into the study. Multiple psychometric and biometric measures are obtained. The range of metrics are collected across a variety of environmental settings, including at different times and during different activities. A fieldwork setting is selected for the study that is relevant to construction, noting that the focus of this study is to demonstrate the importance of ecological validity across fieldwork settings, not just for onsite construction work. Rather than having to deal with the full extremes of a construction site, the fieldwork setting selected is a construction management classroom. Clearly, the construction management classroom is not generally as dynamic as the construction site, but it can be considered directly equivalent to onsite training rooms or to many construction office settings. The overarching characteristic is that it is not a laboratory-based/contrived setting. It is a live, ecologically valid fieldwork setting, where the research instruments have to be straight-forward to administrate, robust, and unobtrusive to the participants.

Furthermore, the study explicitly does not seek to draw conclusions about the particular drivers of stress or how participants respond to stressors in a specific workplace context. It is not a study of stress of itself, but rather a study that seeks to demonstrate how stress can be measured effectively in a construction industry setting. For that reason, the number of study participants is far less significant than the number of metrics and the range of environmental factors included in the study. Indeed, limiting the number of participants to just two, enables a more explicit comparison of results between the participants. Having many more participants would allow for statistical comparisons of results, but the number of participants required to make significant statistical tests across so many variables and contexts would be prohibitive. This purpose and approach adopted for this study only requires two participants, each measured using the same range of metrics, for the periods of time, with the same activities, in the same fieldwork setting.

We present a combination of representative psychometric and biometric measures for two Australian university teachers during a shared, face-to-face, block-teaching assignment, teaching a first-year subject in the Master of Integrated Project Delivery study program at a university in Hong Kong. The teaching block extended over 3 consecutive days (consisting of recorded teaching periods 09:15-16:40, day 1; 08:15-15:35, day, 2; and 09:10-12:35, day 3), to a consistent class cohort of 30 local and mainland Chinese graduate students. Teacher A is a male academic with 40 years teaching experience and a specialisation in construction management. Teacher B is a male academic with, 25 years teaching experience (10 years younger in age than Teacher A) and a specialisation in digital architecture. Neither participant had any known underlying medical conditions that would influence PPG or EDA measurements. Both participants had taught the same subject together, in the same location, on two previous occasions. This was the first occasion the teachers had taught/met this particular cohort of students.

Immediately prior to the teaching assignment, both teachers completed a Perceived Stress Scale (PSS) and State–Trait Inventory of Cognitive and Somatic Anxiety (STICSA) survey instrument to establish the underlying levels of general stress and both the immediate and underlying levels of general anxiety.

The PSS is a 14-item self-report instrument composed of 7 items to measure Perceived Stress and 7 items to measure Perceived Coping. The same instrument can be configured to test responses for a variety of timeframes, from the previous day, to the previous year. In this study the longer term, underlying (trait/tonic) perception of stress over the previous year is used to determine the baseline level of stress for each participant. Each survey item then comprises a statement that always begins with “How often during the last year have you…” and then expresses either a potential stress (“…been upset because of something that happened unexpectedly?”) or a potential coping (“…dealt successfully with irritating life hassles?”). The PPS uses a 5-point scoring scale ranging from “Never (0)”, to “Very often (4)”. The Perceived Coping scores are reversed and summed with the Perceived Stress scores to determine an overall PSS score. This gives a potential overall PSS score in the range of 0 (very low) to 56 (very high) levels of perceived stress.

The STICSA is a, 21-item self-report instrument composed of 10 items to measure cognitive symptoms (thinking) and 11 items to measure somatic symptoms (physical sensations) of general anxiety. The same instrument is used separately to rate how each participant perceives their level of anxiety at that moment in time (the State measure), and in general (the Trait measure). Each item comprises a statement (“I think that others won’t approve of me”, “My muscles are tense”, etc.) on a 4-point ordinal scale. The scale for State and Trait scores range from “not at all/almost never at all (1)” to “very much so/almost always (4)”. A study by Van Dam et al., (2011) indicated that an overall STICSA score of 40 offers an effective cut-off measure for clinical anxiety. Sub-scores of, 23 and 18 were the indicative cut-offs for Cognitive and Somatic measures respectively. Van Dam et al., (2011) also found that the overall Trait score, rather than overall State score, offers the most effective measure of general anxiety.

Prior to each teaching day, a biometric wristband (the Empatica E4: www.empatica.com/research/e4/) was placed on the non-dominant arm of each teacher. The Empatica E4 is equipped with a range of medical-grade sensors selected to gather high frequency, high quality data. It combines PPG and EDA sensors, along with a 3-axis accelerometer to capture movement, an infrared thermopile to record peripheral skin temperature, and a high-accuracy internal real-time clock for efficient synchronisation. The wristband records the raw data from each participant for each session to the local device for later export and analysis. Wearing the biometric wristband all day, each day over a 3-day period, provided a critical workload baseline for the study (Chen and Tserng, 2021).

The fieldwork teaching space utilised a 'horseshoe' style-seating arrangement as an effective way of increasing dynamic interaction between the lecturer and students. A discrete (115mm x 48mm x, 28mm) 360 degree, 5.7K (5760*2880) resolution, 30fps video camera (the Insta360 OneX: www.insta360.com/product/insta360-onex/) was placed to the front, side of the main lecture presentation area, with direct registration of the lecturer presentations, the student reactions, and a wall-mounted time-clock for reference. The camera records in MP4 format and saves directly to an on-board, 256 GB MicroSD Card. Whilst the Insta360 OneX does record sound, to ensure good quality audio rendition, a separate recording was made using a portable high-quality field recorder (the Zoom H4N Pro: https://zoomcorp.com/en/jp/handy-recorders/handheld-recorders/h4n-pro/) using the built-in X/Y stereo microphones. The device supports, 24-bit/96 kHz audio in BWF-compliant WAV format, and saves directly to an on-board 32 GB MicroSD Card. The video and audio recordings ensured precise synchronisation of the activities/behaviour of the participants with their biometric monitoring devices was achieved.

The fieldwork activities comprised a series of different teaching modes/situations across the 3-day teaching block: individual lecture (involving one or other of the two teacher participants presenting to the class); collaborative lecture (involving both participants actively presenting to the class at the same time); student group presentations (presented by each group of 6 students to the rest of the class, and assessed live in-class by both participants); interactive class discussion (between both teacher participants and the entire student cohort); student groupwork (in-class group activities supervised informally by both participants); and meal break (taken by teachers and students communally, inside the teaching space). The two participants were both present in the same teaching space for the entire duration of the study. A breakdown of the timings of the different teaching modes over the entire teaching block, determined from the video/audio recordings, is presented in Figure I.

PLACE FIGURE I HERE

Figure I: The sequence of teaching activities across the 2.5-day block presented in 1-minute intervals

3.2 Data processing

The 14 item PSS survey responses for each participant were divided into Perceived Stress and Perceived Coping subscales following the procedures specified in Ribeiro Santiago et al. (2020). The Perceived Coping score is obtained by reversing the scores (0=4, 1=3, 2=2, etc.) on the seven positive items and summing. Total PSS is then summed across all 14 items. Whilst conversion tables are recommended to adjust total scores for measurement bias when comparing between population groups (Ribeiro Santiago et al., 2020), no adjustment has been made in this instance given identical and very minor adjustments would apply. To provide some context, each score total is also expressed as a percentage of the maximum score possible.

The STICSA survey responses for each participant were separated into state and trait responses, and divided between the cognitive and somatic factors following the procedures specified in Grös et al. (2007). For each factor the scores for relevant items are summed. The total scores for state and trait responses are then simply summed again for comparison purposes.

The Empatica E4 uses a PPG sensor to measure variation in blood volume as it passes to and from the hand with each heartbeat. The BVP is sampled at 64 Hz. Software within the device then identifies peaks in this signal and records realistic temporal distances (more than 0.3 and less than, 2.0s) between these peaks as inter-beat intervals (IBI's). Using a moving 10-s window, HR is then averaged from the inter-beat intervals to remove most of the serious motion artefacts in the raw PPG signal. The resulting 1 Hz (or every second) heart rate mean value is then time-stamped and downloaded as a .csv file for analysis. The raw data is then down-sampled to provide an individual mean HR for each minute of the recordings. Each minute of the recording is then tagged with the appropriate activity (teaching mode) label, derived from the breakdown of the timings of the different teaching modes presented in Figure I. The mean and standard deviation for the HR is then calculated particular to each teaching mode across all periods.

Similarly, the realistic IBI's are individually tagged with the appropriate activity (teaching mode) label, derived from the breakdown of the timings of the different teaching modes presented in Figure I. The collected IBI's for each period and activity are then post-processed to determine the respective PRV using a standard Root Mean Square of Successive Differences (RMSSD) method. The mean for the PRV RMSSD is then calculated particular to each teaching mode across all periods. To render changes in the PRV more apparent, and recognise that individuals may have very different baseline values, only the overall PRV results are presented in absolute values. All other PRV values are presented as a % of that participant baseline value overall.

To demonstrate the potential scale of impact that the circadian cycle can have on biometric measures of stress and anxiety (Scheer et al., 2019), the mean value for the HR and PRV is calculated for combined activities (excluding meal-breaks), separated between AM and PM activities. To investigate the potential scale of impact that familiarity and repetition might have, the mean value for the HR and PRV are also calculated for combined activities (excluding meal-breaks), separated between Day 1, Day, 2 and Day 3. Finally, to compare overall values, the mean for the HR and PRV are calculated for combined activities (excluding meal-breaks), totalled across all 3 days of activities.

The Empatica E4 measures EDA by passing a miniscule electrical current between two electrodes in contact with the skin, and recording the electrical conductance in microSiemens (µS). Raw data from the EDA sensor is sampled at 4Hz, time-stamped and downloaded as a .csv file for analysis. The raw data from the Empatica E4 is then processed in LedaLab (ledalab.de/), a Matlab-based software licensed under the GNU General Public License. The raw data is pre-processed in LedaLab using adaptive smoothing and automatic artefact correction. Continuous Decomposition Analysis (CDA) is then performed to decompose the pre-processed data into continuous signals of phasic and tonic activity, as recommended by Benedek and Kaernbach (2010). The phasic component of each dataset is extracted to indicate event-based arousal. Based on a standard trough-to-peak (TTP) or min-max analysis, LedaLab exports a listing of all significant skin conductance response (SCR) amplitudes (measured in µS) with a corresponding time of onset. Thus, for each significant phasic component, the amplitude (size) of the response and the point in time (nearest second) when the response occurs are both known.

The start and end time (nearest second) for each different teaching mode activity is determined from the sequence of teaching activities presented in Figure I. Using the start/end times as cut-off points, the list of phasic components is blocked into separate periods particular to each teaching mode. The mean amplitude for each block, along with a count of the number of phasic components/events in total for each block, is then calculated. Dividing the total number of events for each block by its duration, the rate of phasic events per minute is determined. The mean amplitude value provides a measure of the relative scale of the SCR for each teaching mode (the size of the response). The rate of events per minute provides a measure of the relative frequency of phasic events for each teaching mode (the frequency of response).

4 Results

A summary of the PSS and STICSA measures that characterise the stress and anxiety levels of the two participants is provided in Table I.

Both participants report relatively low levels of stress overall (Teacher A = 12; Teacher B = 14), considering that the mean overall PSS score for Australian males is, 23-24 (Ribeiro Santiago et al., 2020). This low level of overall PSS is due in large part to the high levels of perceived coping reported by both participants (Teacher A =, 26; Teacher B =, 23). Teacher A has the slightly higher level of perceived stress, but this is offset by a significantly high level of perceived coping. Thus, Teacher B has the slightly higher overall self-reported level of underlying stress (PSS score). However, to some degree this single rating disguises a more complex set of results between the participants.

Both individual teachers generally self-report medium levels of anxiety in total State and total Trait score terms (given the possible range for either score is, 21 to 84 with a clinical cut-off at 40). For Teacher A, the Cognitive to Somatic comparison is similar across both State and Trait scores, although total State anxiety is markedly higher than total Trait anxiety. For Teacher B, the Cognitive to Somatic comparison is almost exactly reversed between the State scores and the Trait scores. Thus, both participants have elevated levels of State Somatic anxiety - close to the indicative cut-off score of 18 (Van Dam et al., 2011). Focussing on the overall Trait score, offering the most effective measure of general anxiety, Teacher B has the slightly higher level of self-reported anxiety, and most particularly in Trait Cognitive terms. Once again, however, the underlying data is more nuanced.

Table I: Psychometric comparison of prevailing teacher self-reported stress and anxiety

PLACE TABLE I HERE

A summary of the mean values determined for the HR, PRV(%) (relative to individual overall baseline), EDA, Amplitude and Rate (frequency) of events for both participants, is provided in Figure II. Results are provided across the entire teaching block (Overall); individually for each day (Day 1, Day, 2, Day 3); split between morning and afternoon (AM, PM); and individually for each teaching mode (Lecture_Self, Lecture_Other, Lecture_Both, Student_Presentation, Student_Interaction, Student_Groupwork, Meal_Break). To summarise results, the ranking of each value in Figure II is shading-coded between the relative best (white), middle (light shading) and worst (dark shading) result values within each metric, across teaching periods and across teaching modes. The results Overall are compared and shading-coded between participants. All other results are compared and shading-coded particular to each participant.

The HR means for each participant are markedly different. Teacher A has a significantly lower/better (shaded white) mean HR overall (84bpm) than Teacher B (110bpm). The lower HR overall for Teacher A is reflected in lower HR for each day, for am and pm, and for each teaching mode. The HR for both participants reduces from Day 1 to Day 3 (87, 83, 80bpm for Teacher A; 118, 107, 99bpm for Teacher B). Where the HR for Teacher A increases slightly from am to pm (84 to 85bpm), the HR for Teacher B falls from am to pm (113 to 107bpm). The highest/worst (shaded dark) HR means for Teacher A are for Lecture_Self (93bpm) and Lecture_Both (88bpm); middle (shaded light) HR means are for Student_Interaction, Student_Groupwork and Meal_Break (84, 84 and 83bpm respectively); and the lowest/best (shaded white) HR means are for Lecture_Other and Student_Presentation (80 and 77bpm respectively). This is in contrast to Teacher B, where the worst HR means are for Lecture_Other (121bpm) and Student_Interaction (118bpm); middle HR means are for Lecture_Both, Student_Presentation, Student_Groupwork and Meal_Break (112, 108, 109 and 109bpm respectively); and the best HR mean is clearly for Lecture_Self (103bpm).

In contrast, it is higher PRV(%) values that generally indicate a more resilient stress response. Thus, Teacher A has a significantly lower (worst) mean absolute PRV overall (0.005) than Teacher B (0.020), which is the exact opposite of the respective HR results. The breakdown and cross-participant comparison of the mean PRV(%) values is then rather variable across periods and activities, and when compared to respective HR values. Where the PVR(%) for Teacher A declines from Day1 to Day 3 (142, 85, 85), the PVR(%) for Teacher B improves over the same period (65, 113, 131). For Teacher A this is in contrast to the equivalent HR relationship, where the HR relationship remains consistent for Teacher B. Where the PVR(%) for Teacher A remains steady from am to pm (at 100), the PVR(%) for Teacher B improves from am to pm (89.20 to 107.70). This largely follows the relative HR values for both participants. Conversely, there is no consistent comparison apparent between the HR and PVR(%) results for the different teaching modes, for either participant. For example, for Lecture_Self, Teacher A goes from worst HR to best PVR(%), whereas Teacher B goes from best HR to middle PVR(%). Contrast that again with Student_Interaction, where there is no change in relative value between HR and PVR(%), for either participant.

The EDA, phasic Amplitude and Rate of events are generally more consistent within each participant, but by no means identical. For the overall means, the relative EDA, Amp and Rate correspond to the relative absolute PVR, and contrasts with the HR for both participants. From Day 1 to Day 3, for Teacher A, the EDA, Amp and Rate are broadly consistent with the HR but contrast with the PVR(%). From Day 1 to Day 3, for Teacher B, the EDA, Amp and Rate contrast with both HR and PVR(%). There is more consistency across all metrics for relative am and pm mean values, but the relative values for Teacher A contrast entirely with Teacher B.

PLACE FIGURE II HERE

Figure II: Comparison of highest, middle and lowest values, shading-coded.

For Lecture_Both and Student_Groupwork activities, between both participants there is general consistency in relative values across all metrics. For Lecture_Self and Student_Interaction activities, between both participants there is general variation in relative values across all metrics. For Lecture_Other and Student_Presentation activities, for Teacher A there is some consistency across all metrics, but for Teacher B more variation. For the Meal_Break activity, this time there is some consistency across all metrics for Teacher B, but for Teacher A some variation.

To further summarise the relative values across all metrics for each period and activity in Figure II, a Total column for each participant has been added and shading-coded. The Total calculation counts 3 points for the best ranking, 2 points for the middle ranking and 1 point for the lowest ranking across each metric. On that basis, the best overall ranking would score 5 x 3 = 15 points, the worst 5 x 1 = 5 points, and the middle rankings would score between 9 and 11 points.

Once again, there are periods and activities where both participants rank equal, slightly different and entirely opposite. According to the Totals calculated in this way, Teacher A becomes less stressful from Day 1 to Day 3, where Teacher B becomes less stressed on Day 2 and more stressed again on Day 3. Where Teacher A becomes considerably more stressed from morning to afternoon, Teacher B becomes very much less stressed over the same period. Both participants are most stressed when both are lecturing at the same time (Lecture_Both). Teacher A is least stressed when Teacher B is lecturing (Lecture_Other) or students are presenting (Student_Presentation). Teacher B is least stressed overall during the Meal_Break. The Total rankings are consistent across both participants for Lecture_Self, Lecture_Both, Student_Interaction, and Student_Groupwork.

5 Discussion

The results of the psychometric instruments administered in this study indicate a close relationship between perceived stress and perceived anxiety. On both key criteria (overall PSS and Trait-Cognitive STICSA) Teacher B self-reports higher levels of stress and anxiety than Teacher A. Notwithstanding, there is sufficient variation in the composition of the stress and anxiety results from both participants to suggest the key criteria have a more nuanced relationship to each other. The key criteria are potentially impacted by: the extent to which higher levels of anxiety compound the level of stress; the extent to which higher levels of stress compound the level of anxiety; the significance of the high level of perceived coping with stress reported by Teacher A, in offsetting the otherwise high level of perceived stress; the potential influence of a higher State anxiety reported by Teacher A; the potential balancing impact of State versus Trait STICSA for both participants; and the potential differences in results if the more generalised stress and anxiety tests were replaced with tests more specific to the stress and anxiety directly associated with teaching (there are tests more specific to public speaking anxiety, for example).

On that basis, neither key criteria can be taken as a surrogate for the other. Separate survey instruments for stress and for anxiety are required if sense is to be made of either. Furthermore, even when the two key criteria provide matching results, as demonstrated in this study, the underlying nature of stress and anxiety creates a complex interaction. Such complex interactions also call for multiple psychometric instruments to be administered to record perceived stress and perceived anxiety across multiple metrics.

The higher levels of perceived stress and anxiety self-reported by Teacher B are supported by the biometric measures of overall HR. The overall mean HR for Teacher B is significantly higher than that for Teacher A. Indeed, the mean HR for Teacher B is consistently higher than that for Teacher A for every teaching period and every teaching activity. If nothing more than the key psychometric criteria and overall mean HR results is considered, then a strong conclusion would be that Teacher B experiences significantly higher stress when teaching than Teacher A.

Contrast that conclusion with the overall mean values measured by PRV, EDA, Phasic Amplitude and/or Phasic Rate. It has been argued that these alternative biometrics provide a more accurate measure of stress than HR, and a more objective measure than the psychometric surveys (Kazar and Comu, 2022). If nothing more than the overall mean values for PRV, EDA, Phasic Amplitude and/or Phasic Rate are considered, then the strong conclusion would be that it is Teacher A who experiences significantly greater stress than Teacher B. Potentially, two utterly contrasting conclusions depending on which metrics are privileged.

Just as the psychometric measures require multiple psychometric instruments across multiple metrics to be considered, so also do the measures of physiology require multiple biometrics to be considered. Indeed, adding further biometric measures such as Cortisol levels and perhaps fNIRS to the analysis would each add to the robustness of any consistent findings. Similarly, including analysis of the tonic component of the EDA and application of alternative decomposition and optimisation protocols to the raw EDA data might also reveal the sensitivity of any conclusions to such choices.

It is also worth noting that these contrasting conclusions only hold in either case, when the mean values overall (across all 3 days of the study) are considered. It is apparent that the conclusions drawn from the results of each individual biometric can change substantially, depending on whether values are taken from Day 1, Day 2, Day 3, am or pm. This study demonstrates that these various time periods can play a significant role in determining stress levels. Similarly, each different activity and each participant introduces a further nuance to how the results could/should be interpreted. Thus, the study also illustrates why the activities of participants is an important factor to incorporate and consider in determining stress levels.

The study results are discussed in the context of how emerging biosensor technologies can be applied most effectively to the study of stress in a construction industry setting. It does not seek, of itself, to draw firm conclusions about the levels or drivers of stress per se. Rather, the purpose is to highlight how variable the results can be using different research instruments, across multiple metrics, with different participants, engaged in different activities, at different times. That is not to say that all results are therefore random or inconclusive. On the contrary, the variability demonstrated promotes the need for multiple metrics, where robust conclusions should only be drawn when results are consistent across those multiple psychometric and biometric measures, and in various environmental settings.

6 Conclusions

Stress can manifest in a variety of forms and is driven by a multitude of factors. It is an incredibly complex condition to study in general (Crosswell and Lockwood, 2020), and most especially difficult to study in the particular fieldwork conditions typical of the construction industry. That complexity needs to be acknowledged and addressed in study designs. Lamentably, numerous studies of stress in construction appear not to grasp the full implications of that complexity. The potential for emerging and accessible biosensor devices to be applied without due consideration for the complexity of the condition, represents a direct challenge to the credibility of such studies. This study demonstrates both the necessity and viability of incorporating multiple psychometric and biometric research instruments in fieldwork studies, across extended periods of time, and for various activities.

The inseparable stress/anxiety complex demands that any study of stress that depends on the measurement of biomarkers be contextualised in both stress and anxiety terms. This study highlights the need for psychometric measures of both perceived stress and perceived anxiety to be included, along with further consideration of how each impacts the other. The particular focus in this study is on the Overall PSS and Trait-Cognitive STICSA. Those generalised instruments might usefully be replaced with survey instruments more specific to the field study activities (such as teaching or public speaking in this study). Further investigation of how perceived coping corelates with perceived stress, and the potential balancing impact of State versus Trait STICSA are both warranted.

The most viable and significant biosensors currently include HR, HRV, EDA, Cortisol levels and fNIRS. Based on the findings of this study, the combined use of multiple sensor types is highly recommended. PRV appears to be a viable surrogate measure for HRV in the context of non-clinical studies of stress. However, all of the biosensors considered require a separate baseline value to be determined for each participant, and that can take hours if not days to establish. Significantly, the use of current consumer-grade EEG devices in a fieldwork study context, is to be discouraged.

Overall, the findings of this demonstration study strongly support the use of multiple biosensors whenever biometric measurements of stress are to be incorporated into a study design. Perceptual, physiological and environmental factors all act in concert to impact stress. Strong conclusions on the potential drivers of stress should only be considered when consistent results across multiple psychometric and biometric research instruments, measuring multiple biomarkers, across various periods of time and a variety of activities, in ecologically valid settings are achieved.

Acknowledgements

The Author wishes to acknowledge the support and participation of the Program Director, Students and Teachers involved in this study.

7 References

Ali, N. and Nater, U.M. (2020), Salivary Alpha-Amylase as a Biomarker of Stress, *Behavioral Medicine. Int. J. Behav. Med.*, Vol.27, pp.337–342.

Benedek, M. and Kaernbach, C. (2010), A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, Vol.190, pp.80–91.

Bystritsky, A. and Kronemyer, D. (2014), Stress and Anxiety: Counterpart Elements of the Stress/Anxiety Complex, *Psychiatr Clin N Am*, Vol.37, pp.489–518.

Campbell, F. (2006), *Occupational Stress in the Construction Industry*, Chartered Institute of Building, Berkshire, UK.

Chae, J., Hwang, S., Seo, W. and Kang, Y. (2021), Relationship between rework of engineering drawing tasks and stress level measured from physiological signals, *Automation in Construction*, Vol.124, pp.103560-1–12.

Chalmers, T., Hickey, B.A., Newton, P., Lin, C.-T., Sibbritt, D., McLachlan, C.S., Clifton-Bligh, R., Morley, J. and Lal, S. (2022), StressWatch: The Use of Heart Rate and Heart Rate Variability to Detect Stress: A Pilot Study Using Smart Watch Wearables, *Sensors*, Vol.22, No.151, doi:10.3390/s22010151.

Chan, A.P.C., Nwaogu, J.M. and Naslund, J.A. (2020), Mental Ill-Health Risk Factors in the Construction Industry: Systematic Review, *J. Constr. Eng. Manage.*, Vol.146, No.3, pp.04020004-1–13.

Chen, W-C. and Tserng, H.P. (2021), Real-time individual workload management at tunnel worksite using wearable heart rate measurement devices, *Automation in Construction*, Vol.134, pp.104051-1–15.

Choi, B., Jebelli, H. and Lee, SH. (2019), Feasibility analysis of electrodermal activity (EDA) acquired from wearable sensors to assess construction workers' perceived risk, *Safety Science*, Vol.115, pp.110–120.

Cohen, S., Gianaros, P. and Manuck, S. (2016), A stage model of stress and disease, *Perspectives on Psychological Science*, Vol.11, No.4, pp.456–463.

Crosswell, A.D. and Lockwood, K.G. (2020), Best practices for stress measurement: How to measure psychological stress in health research, *Health Psychology Open*, July-December, pp.1–12.

Epel, E.S., Crosswell, A.D., Mayera, S.E., Prathera, A.A., Slavichb, G.M., Putermanc, E. and Mendesa, W.B. (2018), More than a feeling: A unified view of stress measurement for population science, *Frontiers in Neuroendocrinology*, Vol.49, pp.146–169.

Fink, G. (2017), Stress: Concepts, Definition and History, in G. Fink, *Reference Module in Neuroscience and Biobehavioral Psychology*, Elsevier, pp.1–9.

Geršak, G. (2020), Electrodermal activity - a beginner's guide, *Electrotechnical Review*, January, pp.175–182.

Grös, D.F., Antony, M.M., Simms, L.J. and McCabe, R.E. (2007), Psychometric Properties of the State–Trait Inventory for Cognitive and Somatic Anxiety (STICSA): Comparison to the State–Trait Anxiety Inventory (STAI), *Psychological Assessment*, Vol.19, No.4, pp.369–381.

Hagermoser Sanetti, L.M., Boyle, A.M., Magrath, E., Cascio, A. and Moore, E. (2020), Intervening to Decrease Teacher Stress: a Review of Current Research and New Directions, *Contemporary School Psychology*, https://doi.org/10.1007/s40688-020-00285-x.

Hwang, S., Seo, JO., Jebelli, H. and Lee, SH. (2016), Feasibility analysis of heart rate monitoring of construction workers using a photoplethysmography (PPG) sensor embedded in a wristband-type activity tracker, *Automation in Construction*, Vol.71, pp.372–381.

Järvelin-Pasanen, S., Sinikallio, S. and Tarvainen, M.P. (2018), Heart rate variability and occupational stress – systematic review, *Industrial Health*, Vol.56, pp.500–511.

Jebelli, H., Choi, B. and Lee, S.H. (2019), Application of Wearable Biosensors to Construction Sites. I: Assessing Workers' Stress, *J. Constr. Eng. Manage.*, Vol.145, No.12, pp.04019079-1–12.

Jebelli, H., Hwang, S. and Lee, SH. (2018), EEG-based workers' stress recognition at construction sites, *Automation in Construction*, Vol.93, pp.315–324.

Jeon, JH. and Cai, H. (2021), Classification of construction hazard-related perceptions using: Wearable electroencephalogram and virtual reality, *Automation in Construction*, Vol.132, doi:10.1016/103975.

Jiang, X., Bian, G-B. and Tian, Z. (2019), Removal of Artifacts from EEG Signals: A Review, *Sensors*, Vol.19, No.987, doi:10.3390/s19050987.

Julian, L.J. (2011), Measures of Anxiety, *Arthritis Care and Research*, Vol.63, No.S11, pp.S467–S472.

Kamardeen, I. (2022), Work stress related cardiovascular diseases among construction professionals, *Built Environment Project and Asset Management*, Vol.12, No.2, pp.223–242.

Karatsoreos, I.N. (2018), Stress: Common themes toward the next frontier, *Frontiers in Neuroendocrinology*, Vol.49, pp.3–7.

Kazar, G. and Comu, S. (2022), Exploring the relations between the physiological factors and the likelihood of accidents on construction sites, *Engineering, Construction and Architectural Management*, Vol.29, No.1, pp.456–475.

Ke, J., Du, J. and Luo, X. (2021), The effect of noise content and level on cognitive performance measured by electroencephalography (EEG), *Automation in Construction*, Vol.130, pp.103836-1–14.

Kim, H-G., Cheon, E-J., Bai, D-S., Lee, Y.H. and Koo, B-H. (2018), Stress and Heart Rate Variability: A Meta-Analysis and Review of the Literature, *Psychiatry Investig*, Vol.15, No.3, pp.235–245.

Kreibig, S.D. and Gendolla, G.H.E. (2014), Autonomic Nervous System Measurement of Emotion in Education and Achievement Settings, in R. Pekrun and L. Linnenbrink-Garcia, (Eds.) *International Handbook of Emotions in Education*. ProQuest Ebook Central, pp.625–642.

Langdon, R.R. and Sawang, S. (2018), Construction Workers' Well-Being: What Leads to Depression, Anxiety, and Stress? *J. Constr. Eng. Manage.*, Vol.144, No.2, pp.04017100-1–15.

Lau, T.M., Gwin, J.T. and Ferris, D.P. (2012), How Many Electrodes Are Really Needed for EEG-Based Mobile Brain Imaging?, *Journal of Behavioral and Brain Science*, Vol.2, pp.387–393.

Lazarus, R. S. and Folkman, S. (1987), Transactional theory and research on emotions and coping, *Eur. J. Personality*, Vol.1, No.3, pp.141–169.

Lee, E-H. (2012), Review of the Psychometric Evidence of the Perceived Stress Scale, *Asian Nursing Research*, Vol.6, pp.121–127.

Lee, W., Lin, K-Y., Seto, E. and Migliaccio, G.C. (2017), Wearable sensors for monitoring on-duty and off-duty worker physiological status and activities in construction, *Automation in Construction*, Vol.83, pp.341–353.

Love, P. E., Edwards, D. J. and Irani, Z. (2010), Work stress, support, and mental health in construction. *J. Constr. Eng. Manage.*, Vol.136, No.6, pp.650–658.

Mejía-Mejía, E., Budidha, K., Abay, T.Y., May, J.M. and Kyriacou, P.A. (2020a), Heart Rate Variability (HRV) and Pulse Rate Variability (PRV) for the Assessment of Autonomic Responses, *Frontiers in Physiology*, Vol.11, No.779, doi:10.3389/fphys.2020.00779.

Mejía-Mejía, E., May, J.M., Torres, R. and Kyriacou, P.A. (2020b), Pulse rate variability in cardiovascular health: a review on its applications and relationship with heart rate variability, *Physiological Measurement*, Vol.41, doi:10.1088/1361-6579/ab998c.

Nwaogu, J.M. and Chan, A.P.C. (2021), Work-related stress, psychophysiological strain, and recovery among on-site construction personnel, *Automation in Construction*, Vol.125, pp.103629-1–13.

Quaresima, V. and Ferrari, M. (2019), Functional Near-Infrared Spectroscopy (fNIRS) for Assessing Cerebral Cortex Function During Human Behavior in Natural/Social Situations: A Concise Review, *Organizational Research Methods*, Vol.22, No.1, pp.46–68.

Ribeiro Santiago, P.H., Nielsen, T., Smithers, L.G., Roberts, R. and Jamieson, L. (2020), Measuring stress in Australia: validation of the perceived stress scale (PSS-14) in a national sample, *Health and Quality of Life Outcomes*, Vol.18, No.100, https://doi.org/10.1186/s12955-020-01343-x.

Saedi, S., Fini, A.A.F., Khanzadi, M., Wong, J., Sheikhkhoshkar, M. and Banaei, M. (2022), Applications of electroencephalography in construction, *Automation in Construction*, Vol.133, pp.103985-1–15.

Samson, C. and Koh, A. (2020), Stress Monitoring and Recent Advancements in Wearable Biosensors, *Front. Bioeng. Biotechnol.*, Vol.8, No.1037, doi: 10.3389/fbioe.2020.01037.

Scheer, F.A.J.L., Chellappa, S.L., Hu, K. and Shea, S.A. (2019), Impact of mental stress, the circadian system and their interaction on human cardiovascular function, *Psychoneuroendocrinology*, May, No.103, pp.125–129.

Styck, K.M., Rodriguez, M.C. and Yi, E.H. (2020), Dimensionality of the State–Trait Inventory of Cognitive and Somatic Anxiety, *Assessment*, Vol.29, No.2, pp.103–127.

Textor, M. (2019), Perceptual objectivity and the limits of perception, *Phenom Cogn Sci*, Vol.18, pp.879–892.

Van Dam, N.T., Gros, D.F., Earleywine, M. and Antony, M.M. (2011), Establishing a trait anxiety threshold that signals likelihood of anxiety disorders, *Anxiety, Stress and Coping*, Vol.26, No.1, pp.70–86.

Wexler, A. and Thibault, R. (2019), Mind-Reading or Misleading? Assessing Direct-to-Consumer Electroencephalography (EEG) Devices Marketed for Wellness and Their Ethical and Regulatory Implications, *Journal of Cognitive Enhancement*, Vol.3, 131–137.

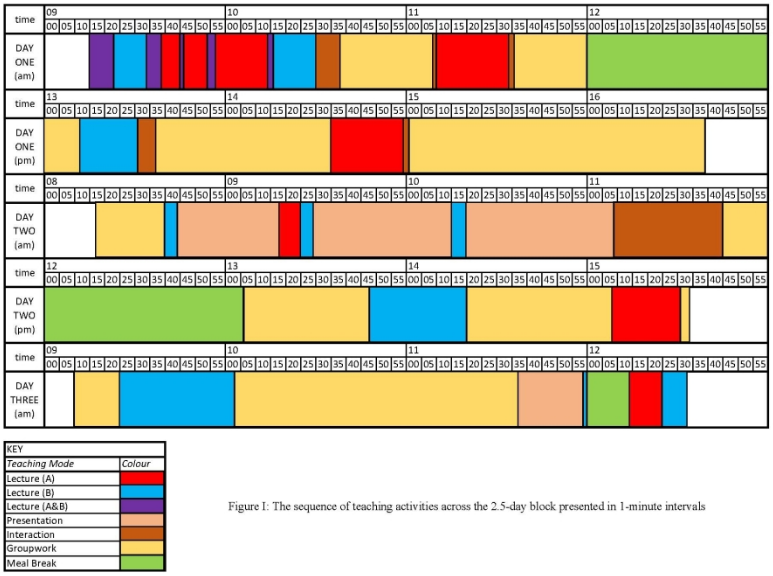

Figure I: The sequence of teaching activities across the 2.5-day block presented in 1-minute intervals

297x210mm (200 x 200 DPI)

The study of stress and anxiety in construction: a case for the use of a multimodal approach in fieldwork settings

Table I: Psychometric comparison of prevailing teacher self-reported stress and anxiety

Stress Construct	**Teacher A**	**Teacher B**
Perceived Stress	10	9
relative to maximum	*35.70%*	*32.10%*
Perceived Coping	26	23
relative to maximum	*92.86%*	*82.14%*
Overall PSS	12	14
relative to maximum	*21.43%*	*25.00%*
Anxiety Construct	**Teacher A**	**Teacher B**
State-Cognitive	14	13
State-Somatic	17	16
Total State	*31*	*29*
Trait-Cognitive	11	16
Trait-Somatic	14	12
Total Trait	*25*	*28*

Teaching Period	Teacher A						Teacher B					
	HR	PRV(%)	EDA	Amp	Rate	TOTAL	HR	PRV(%)	EDA	Amp	Rate	TOTAL
Overall	84.27	0.005	1.18	0.27	21.42		110.26	0.020	1.00	0.13	6.60	
Day 1	86.97	142	1.32	0.34	34.23	**6**	118.42	64	1.15	0.12	6.03	**9**
Day 2	83.47	85	1.15	0.21	12.49	**11**	107.34	113	0.61	0.11	6.50	**12**
Day 3	80.08	85	0.95	0.24	12.94	**11**	98.74	131	1.54	0.17	7.93	**9**
AM	83.89	100	1.17	0.26	19.85	**13**	112.86	89	1.26	0.14	6.44	**5**
PM	84.66	100	1.19	0.26	22.09	**7**	107.52	108	0.74	0.09	6.29	**15**
Lecture_Self	93.01	145	1.53	0.29	27.79	**10**	103.40	92	1.31	0.13	9.37	**9**
Lecture_Other	79.79	113	1.00	0.26	17.64	**12**	121.10	121	1.20	0.09	4.34	**11**
Lecture_Both	88.41	100	1.40	0.40	30.21	**6**	111.78	52	1.57	0.18	8.93	**6**
Student_Presentation	76.83	109	0.97	0.24	13.92	**12**	108.47	125	1.13	0.15	5.95	**9**
Student_Interaction	83.92	106	1.66	0.24	20.80	**10**	118.47	45	0.60	0.09	5.05	**10**
Student_Groupwork	84.29	74	1.20	0.27	22.80	**9**	109.11	90	1.06	0.13	6.68	**10**
Meal_Break	83.00	174	0.82	0.37	22.37	**11**	108.99	116	0.33	0.04	3.81	**14**

Key		
Value	Best	3-point
Value	Middle	2-points
Value	Worst	1-points

Figure II: Comparison of highest, middle and lowest values, shading-coded.

536x305mm (96 x 96 DPI)