CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	From the lab to the living room: examining challenges with the transition of heart rate variability measurement to real world contexts
Authors	Jump, Owen
Publication date	2021-07-01
Original Citation	Jump, O. 2021. From the lab to the living room: examining challenges with the transition of heart rate variability measurement to real world contexts. PhD Thesis, University College Cork.
Type of publication	Doctoral thesis
Rights	© 2021, Owen Jump. - https://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2026-09-16 18:27:38
Item downloaded from	https://hdl.handle.net/10468/13279

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

From the lab to the living room: Examining challenges with the transition of heart rate variability measurement to real world contexts

Owen Jump, School of Applied Psychology, University College Cork.

This is a publication-based thesis submitted for the degree of PhD by research to National University of Ireland, Cork
School of Applied Psychology

Submission date: July 1st 2021

Head of School: Prof Carol Linehan

Primary supervisor: Dr Samantha Dockray

Secondary supervisor: Dr Mike Murphy

Table of contents

List of Tables……6

List of Figures……..7

List of Acronyms ……8

List of Outputs ……9

Declaration of Academic Honesty ……...10

Acknowledgements ……...11

Abstract……..13

Chapter 1 Introduction……15

1.1 General Introduction…….15

1.2 Heart rate measurement……23

1.3 Theoretical frameworks……34

1.4 Laboratory testing of Heart Rate Variability……49

1.5 Ambulatory measurement……52

1.6 Experience sampling……55

Chapter 2: Measuring the psychobiological correlates of daily experience in adolescents ……....61

2.1 Abstract……61

2.2 Introduction…….....62

2.3 Ecological Momentary Assessment……63

2.4 Measuring Salivary Analytes in Ambulatory Settings……64

2.5 Sampling Protocols……..67

2.6 Saliva Collection Devices……....68

2.7 Supporting Participant Adherence……...69

2.8 Sampling Kits……..71

2.9 Measuring Cardiovascular Activity……72

2.10 Measuring Heart Rate in Ambulatory Contexts……....73

2.11 Practical considerations for HR Measurement……..75

2.12 Measuring Sleep……....76

2.13 Objective and Subjective Measures of Sleep……....78

2.14 New Technologies for Psychobiological Measurements……..80

2.15 Cardiovascular Measurement Devices and Application……...80

2.16 Validity of Commercial Applications……...81

2.17 Recording experience and affect……...83

2.18 New horizons……....84

Chapter 3 Methodology……...86

3.1 Methodology overview……86

3.2 Participants……..89

3.3 Materials and Measures……...90

3.4 Study Procedures……103

3.5 Data Analysis……..112

3.6 Pilot Phase……....117

3.7 Ethics……....119

Chapter 4: Study two - Cardiovascular responses to stress utilising anticipatory singing tasks ……120

4.1 Abstract……...120

4.2 Introduction…….122

4.3 Method……127

4.4 Findings……..136

4.5 Discussion……...141

Chapter 5: Study three- Examining the use of an online adaptation of the Day Reconstruction Method ……145

5.1 Abstract……..145

5.2 Introduction……146

5.3 Method……...152

5.4 Findings…….156

5.5 Discussion…….170

Chapter 6: Study Four - From the lab to the living room: Measuring heart rate variability in ecologically valid contexts ……178

6.1 Abstract……..178

6.2 Introduction……180

6.3 Method……...187

6.4 Findings……..198

6.5 Discussion……..209

Chapter 7: Discussion …….214

7.1 General Discussion……..214

7.2 Laboratory measurement of HRV and affect……...215

7.3 Capturing experience and affect inside and outside of the laboratory…….219

7.4 Matching laboratory and ambulatory contexts using affect and experience…224

7.5 The relationship of affect and laboratory and ambulatory HRV measures…..227

7.6 Conclusions……..230

References ……235

Appendices ……..283

List of tables

Table 1.1: *Time domain measures of HRV……29*

Table 1.2: *Frequency domain measures of HRV……30*

Table 1.3*: Principles of social engagement……40*

Table 3.1: *Range of Actiheart HRV Measures……93*

Table 3.2: *A sample of checklist for context for positive affect ambulatory HRV……91*

Table 4.1. *Descriptive Data for Resting, Reactivity and Recovery Conditions……136*

Table 5.1: *Total episodes reported by Activity, Social interaction and social interaction...157*

Table 5.2*: Mean affect rating by episode type*……..160

Table 5.3: *Mean affect rating by social interaction*……..159

Table 5.4: *Mean affect rating by time of day*……....163

Table 6.1: *A sample of checklist for affect context and ambulatory HRV*……....196

Table 6.2: *Overall Mean ambulatory affect rating per episode type*……...199

Table 6.3: *Mean affect rating by laboratory condition*……200

Table 6.4. *Analysis of Variance for Resting, Reactivity, Recovery Laboratory and Ambulatory Positive affect conditions*……..202

Table 6.5: *Multilevel model for BPM, RMSSD and High frequency measures of HRV in laboratory and ambulatory contexts*…….206

List of figures

Figure 1.1: *The PQRST trace*……24

Figure 1.2: *Multilevel NVI structure*…….44

Figure 3.1: *Actiheart signal test user interface*…….95

Figure 3.2*: Sample signal test report*……96

Figure 3.3: *Overview diagram of SSST timeline and phases*……98

Figure 4.1: *Diagram of SSST timeline and phases*……..131

Figure 4.2: *Mean BPM for Baseline,*
Maths, Singing and Recovery conditions…….137

Figure 4.3: *Mean IBI for Baseline, Maths,*
Singing and Recovery conditions…….137

Figure 4.4: *Mean RMSSD for Baseline,*
Maths, Singing and Recovery conditions……5……...137

Figure 4.5 *Mean HF for Baseline, Maths,*
Singing and Recovery conditions…….137

Figure 5.1: *Time of day by positive affect rating*……....165

Figure 5.2: *Time of day by negative affect rating*……...166

Figure 5.3: *Time of day by tiredness rating*……....167

Figure 5.4*: Daily adherence percentages,*
categorised as Full, Part and Non-completion……....169

List of acronyms

ANOVA – Analysis of Variance
ACTH – Adrenocorticotropic Hormone
BP – Blood Pressure
BPM – Beats per Minute
CAR – Cortisol Awakening Response
CV – Cardiovascular
DHEA – Dehydroepiandrosterone
DRM – Day Reconstruction Method
DRM-E – Day Reconstruction Method Electronic
ECG – Electrocardiogram
EMA – Ecological Momentary Assessment
ESM – Experience Sampling Methodology
GPS – Global Positioning System
HF – High Frequency
HPA Hypothalamic Pituitary Adrenal (Axis)
HR – Heart Rate
HRV – Heart Rate Variability
IBI – Inter Beat Interval
LF – Low Frequency
LF/HF – Low Frequency/High Frequency (Ratio)
NVI – Neurovisceral Integration (Model)
PA- Physical Activity
pNN50% - Percentage of successive RR intervals that differ by more than 50 ms
QR- Quick Response (Codes)
REM – Rapid Eye Movement (Sleep)
RMSSSD-Root Mean Square of Successive RR interval differences
RRR- Resting/Reactivity/Recovery
SD – Standard Deviation
SDRR- Standard deviation of RR intervals
SDANN- Standard deviation of the average NN intervals
SDNN - Standard deviation NN Intervals
SPSS – Statistical Package for the Social Sciences
SSST – Sing a Song Stress Task
TSST – Trier Social Stress Task
TOD – Time of Day
ULF - Ultra Low Frequency
VLF – Very Low Frequency

LIST OF OUTPUTS

Jump, O., & Dockray, S. (2021) From the lab to the living room: Measuring heart rate variability in ecologically valid contexts. Journal of psychophysiology (Under review)

Jump, O., & Dockray, S. (2020) Cardiovascular responses to stress utilizing anticipatory singing tasks. Journal of Psychophysiology (Preprint) https://doi.org/10.1027/0269-8803/a000269

Jump, O., & Dockray, S. (2020) From the Lab to the Living Room: Measuring Heart Rate Variability in Ecologically Valid contexts: Psychological Society of Ireland, Division of Health Psychology Annual Conference

Dockray, S., O'Neill, S., & Jump, O. (2019). Measuring the psychobiological correlates of daily experience in adolescents. Journal of Research on Adolescence, 29(3), 595-612. https://doi.org/10.1111/jora.12473

Jump, O., & Dockray, S (2017) Social interactions and heart rate variability: From the lab to the living room: Psychological Society of Ireland, Division of Health Psychology Annual Conference.

Jump, O., & Dockray, S (2017) A variant of the Trier Social Stress Task using Sing-A-Song and computer-based tasks Psychological Society of Ireland, Division of Health Psychology Annual Conference.

Declaration of academic honesty.

I hereby declare that this work is my own and has not been submitted elsewhere or previously for a degree anywhere. Where the work of other is mentioned or quoted it has been appropriately referenced.

Signed

_____

Owen Jump

Acknowledgements

This work was only completed because of the people who have been with me on this journey. Although my name is on the front but there should be many more beside it.

For my family,

Nollaig, you are the reason this PhD happened. Without your love and patience, we would not have finished it. How you put up with me through his process I'll never know! Who you are as a mother, a teacher, and a person is everything I aspire to and this PhD is for you above anyone else.

Fionn, you are not old enough to know yet how much you helped me but being a dad to you got me through the toughest days. You have made me laugh since the day you were born and inspired me to continue. I hope someday you will read this and know you were with me on the journey.

Mam, you are the reason I went back to education, how you have faced the adversity in your life has inspired me to this day. Thank you for all your support.

Finn and Granda, you supported me the way most Irish people show love, by feeding me, I would literally have starved in the early days without you!

Peter, thanks for being there to listen and support me on our many surf trips, your strength as brother helped me to face back into this challenge many times.

For my supervisors,

Samantha, your dedication, and commitment to your teaching, research, and the personal development of your students and the betterment of everyone around you is an inspiration. I have seen you give so much to those around you, including me, and I am honoured to have you

as one of the many mentors that have been part of my life. Thank you for all of your time and support over these past years.

Mike, although we didn't work as closely, you have always been available and supported me whenever I asked, thank you for that.

To my colleagues,

in CIRTL, you have helped me in so many ways in the past couple of years and I am so grateful for your time and patience, in particular Catherine who, like Samantha has shown me the joy in helping others.

And finally, my fellow PhD students and now friends. If it takes a village to raise a child, it takes a room that looks like a fallout shelter with bare concrete blocks for walls full of people on a common journey to raise a PhD student. The laughter and friendship I found in that room will be my longest standing memory of my time as a postgraduate student.

Thank you all sincerely

Abstract

Cardiovascular activity has been widely incorporated as a measure in psychophysiological science in laboratory contexts and has been associated with a variety of health outcomes. In order to study patterns of heart rate in laboratory contexts, researchers have measured tonic/resting, and phasic/reactive Heart Rate Variability, hereafter HRV, and advances in technology enables prolonged periods of data capture. Although the quality of these data is still disputed, the transition to ambulatory measurement is accelerating and research that links ambulatory and laboratory contexts is timely. The PhD presented here describes challenges in laboratory and ambulatory measurement, including heart rate measurement and experience sampling considerations that arise in studies using comparative data from the laboratory and ecologically valid real-world contexts. Four studies are presented that combine as a body of work examining this area of enquiry, initially by detailing measurement of heart rate and how it responds to challenge in the laboratory and then how these measures extend to ecologically valid contexts.

Study one (Chapter two) is a summary narrative review outlining the practical challenges and applications of ambulatory measurement of psychophysiological measures for adolescents, a cohort for whom different challenges in measurement may exist. The review encompasses ambulatory measurement, including sampling techniques heart rate, details practical approaches, and outlines questions for utilising available technologies, including the pragmatic aspects of collecting ambulatory data that are salient in working with adolescent participants.

Study two (Chapter four) details a novel laboratory stressor paradigm that integrates technical adjustments to Electrocardiogram (ECG) measurement required for Heart Rate Variability (HRV) analysis. The study demonstrates the task efficacy, including an examination of anticipatory stress, and proposes its suitability as a potential replacement constituent or complementary task component for the Trier Social Stress Task (TSST). We suggest that this

task can be considered where participants are required to have multiple interactions with the TSST, and where external habituation effects, such as rehearsal effects.

Study three (Chapter five) presents an adaptation of the Day Reconstruction Method (DRM) and demonstrates a novel means of assessing participant adherence to daily diary measures. It details observed patterns of participant adherence and implications for studies using DRM instrumentation, including the ability to assess data with a granularity that was not possible previously. The study details how the integration of retrospective and momentary sampling is desirable, in terms of the methodological benefits, the quality of data collected, and how this better reflects the concept of measurement surrounding the remembered and experiencing self, commenting on how this dichotomy should be resolved to better capture human experience.

Study four (Chapter six) details how HRV measurement in ambulatory contexts, laboratory measures, and experience sampling are combined in the major study for this PhD. Study four extends the measurement of participants' HRV patterns from laboratory contexts, and then using HRV measures and experience sampling, examines how they relate to ambulatory patterns captured across four full days of 24 hour-a-day measurement for each participant. This study examines patterns of baseline and stressor HRV inside the laboratory, extends to measure matched ambulatory measurement rest periods, and finally laboratory reactivity patterns and ambulatory baselines. Multilevel modelling is used to describe group and fixed effects and how participants' patterns of HRV change or remain similar as they move across conditions. The study demonstrates a novel means of screening ambulatory data, using affective patterns and experience sampling to associate it with laboratory measures for HRV type studies.

The combined work presented provides novel insights into laboratory and ambulatory measurement of HRV and provides evidence for the linking measurement in both contexts.

Chapter 1: Introduction

1.1 General introduction

The story of heart rate measurement in humans *is* the story of the technology available to measure it at any given time. The first reliable evidence of heart rate as implicated in biological function was documented by Galen (131 -200 AD). Galen first described the timing of the heart beat as "Pulse", and these descriptions dominated medical science for almost 16 centuries until the early 18th century when the first major revolution in measurement of the heart was brought about by improvements in technology that uncovered variations in blood pressure, owing to exertions (Billman, 2011). The earliest animal and human demonstrations were famously brutal, including Stephan Hales' glass tube experiment, which demonstrated the immense pressures exerted inside a horses' circulatory system by attaching a transparent tube that allowed a visualisation of the dynamic pressures of the heart. Advances in the accuracy of chronological devices further facilitated the development of hales work and beat to beat cycles were described from arterial pressure measurements. Devices such as Floyers' "physician pulse watch" (Floyer,1710) allowed heart rate and respiration to be catalogued. In 1845 Carl Ludwig first observed respiratory sinus arrhythmia (RSA), using a "drum kymograph" in a series of experiments on dogs (Ludwig, 1847). The first electrical imprint of the heart was measured by Wilhelm Einthoven with the development of galvanometers (Barold, 2003) and this led to the ability to determine beat-to-beat patterns in cardiac rhythm. With advances in technology and the realisation that beat-to-beat variations could be categorised into various time and frequency domains scientists began the process of describing and delineating these signals in precise detail.

As the influence of respiration and neural inputs were proposed it was soon realised that heart rate was also subject to mental processes and emotional arousal, and it followed that these data would be associated with psychological indices and subsequently begun the interests of

Psychophysiologists. By the early twentieth century Langley (along with others) had proposed the term “autonomic function” and begun to describe the how vagal and sympathetic systems operated (Berntson & Cacioppo, 2014). As with most of psychology, as the science was advancing, it moved into the laboratory with the behaviourists and researchers became more and more concerned with causality, favouring the identification of a unitary process that linked to behaviour. This led to the pursuit of isolated biological structures, and experiments being conducted in tightly controlled conditions. This tendency to examine biology in alignment with behaviour, including heart rate, led to the belief that humans were merely the summation of classically conditioned nervous system responses, a type of grand stimulus-response machine, that once quantified and accounted for with the correct methodologies would combine to explain the whole in a sort of unified theory of biology and psychology. We know now that this simply is not the case and although the physiological correlates of Heart Rate (HR) measurement are still contested (Billman, 2011), the field of psychophysiology has made considerable advances in describing the myriad psychological processes that link to, and indeed shape, the outputs generated by the heart. Theory relating to biology and psychological function now views the nervous system as a series of integrated systems (Thayer et al., 2012) that have afferent, efferent and interconnected components that continuously feedback to central and peripheral regions, providing the wonder that is the homeostatic balance, that have allowed us not to not only to exist in our environment - but to adapt and thrive.

As the development of small portable devices such as the Holter ™ monitor (Holter, 1961) made ambulatory measurement possible, physiological measurement could, for the first time, move outside of the laboratory into these environments. This highlighted the ever-present question in psychophysiology; how do we accurately measure naturalistic processes while retaining the precision and accuracy afforded by laboratory-based measures? Do we err on the

side of caution and rigorously control almost every conceivable variable and risk losing the ability to generalise or claim to ecological validity? Or do we allow participants to roam "in the wild", in uncontrolled environments and risk losing control of our data to the point that it becomes so noisy with everyday activity and physiological confounds, such as movement, that it becomes useless? The central problem we must solve in psychophysiology is ecological validity and Stephan Porges summarised this challenge stating,

A recognized fact which goes back to the earliest times is that every living organism is not the sum of a multitude of unitary processes, but is, by virtue of interrelationships and of higher and lower levels of control, an unbroken unity. When research, in the efforts of bringing understanding, as a rule examines isolated processes and studies them, these must of necessity be removed from their context (Porges, 2007, p301)

Now as the advancement of technology has facilitated new methods of measurement it has become even more important for us to re-examine this question. Wearable technologies and free mobile applications are now widely available, and while previously ambulatory technologies had been a reality, they were often clunky, awkward, and of questionable accuracy, particularly when regarded in the context of research. However, wrist watches now carry ECG capability, discreet chest straps can send uninterrupted data streams to smartphones, and a myriad of 'health' applications allow people to track their own cardiovascular systems in real time. While uncertainty exists surrounding the accuracy of many of these devices (Hernando et al., 2018; Malik et al., 2018), and this area is under researched, it is incontrovertible that they are already being used across ever widening contexts. Companies such as Apple™, Fitbit™, and Samsung™ now hold vast repositories of recorded heartbeats (Reiss et al., 2019) with other companies, although not to the same extent, holding similar data

sets. While the quality of these data can be debated with limited capability on many devices, one fact cannot be ignored; they exist and they are already being used, both by the individuals that wear the devices and the companies that have a proprietary claim to the data.

Although the physiological data returned by these devices are currently limited, for example, most smart watches use photoplethysmography that only capture vascular dilation of larger blood vessels at the wrist, their accuracy is improving with each iteration with some now capable of ECG measurement. Soon researchers may have to consider the use of these devices as equivalent and perhaps more convenient and affordable than expensive clinical grade devices. Certainly, in studies where capturing the biological data is not the primary objective there are already incentives to use these more easily obtained data. While this transition might not necessarily be imminent, it is happening, and this represents a timely opportunity for psychophysiology because we are in a fortunate position to contribute to the transition of these measurement techniques as they are integrated into people's daily lives.

The future study of ambulatory measurement of HRV and its psychosocial correlates will answer a number of key challenges Firstly, how prolonged exposure to specific situationally located stressors are linked to health (Smith, Deits-Lebehn, Williams, Baucom, & Uchino, 2020). Future study should have as its primary aim the improvement of health, both physical and psychological with end user data that is accessible and informative. Secondly, the generally accepted hypothesis that laboratory measurement is related to everyday life will be more widely examined and questions such as those related to behavioural mediation and its impact on exposure to stressors will be answered. (McEwen, 2001), for example, how common is the exposure to a serial supervised maths task in everyday life? Thirdly, as technological advances allow more people to measure and interpret their own physiological data research is required

to contextualise and authenticate these data and validate or dismiss them in relation to laboratory measures (Brown et al., 2017), this will include statistical modelling techniques to account for confounds such as movement, for example using accelerometer data, to allow situationally relevant HRV measurement (Laborde et al., 2017). The ultimate goal being the integration and appropriate use of ambulatory and laboratory study that are not in opposition but complimentary.

The strength of our science is the scientific processes and standards of measurement we employ, and we can leverage this to contribute to research outside of the laboratory by ensuring that best practice is adhered to. Within psychology there is a legacy of proven methods that have been demonstrated in the laboratory, and this knowledge can support proponents of ambulatory technologies to study the transition between the laboratory and the living room, ensuring that structured and properly thought-out practices are adhered to. It is important then that research, whether carried out in the laboratory or ambulatory contexts, be grounded in this science. This is the focus of the PhD presented here, to contribute to the transition from the laboratory to the living room. This dissertation describes some key challenges in ambulatory measurement, including heart rate measurement and experience sampling, that arise in studies using comparative data from the laboratory and ecologically valid real-world contexts. Four studies are presented that combine as a body of work examining laboratory and ambulatory measurement that contributes to the study of psychophysiological measurement, initially by detailing measurement of heart rate in the laboratory and then examining how these measures, anchored in affect and experience, extend to ecologically valid contexts.

Chapter two presents study one, “Measuring the psychobiological correlates of daily experience in adolescents” (Dockray, O’Neill & Jump, 2019). This is a summary narrative

review outlining the application of ambulatory measurement for a wide range of psychophysiological measurement but with reference to a specific cohort, namely adolescents. The review encompasses ambulatory measurement, including sampling techniques for cortisol collection, with special consideration for working with adolescent participants. The study also outlines questions for utilising available technologies and mobile applications, measurement of heart rate in ambulatory contexts, and finally experience sampling methodologies. This study has been published in the Journal of Research on Adolescence and is detailed in chapter three.

Chapter four details study two, "Cardiovascular Responses to anticipatory stress utilising anticipatory singing tasks" (Jump & Dockray, 2020). This study is an original laboratory based experimental task examining an adaptation of a stressor paradigm proposed by Brouwer (Brouwer & Hogervorst, 2014), that integrates the technical adjustments, both in terms of ECG measurement and methodological requirements for Heart Rate Variability (HRV) analysis. Study two demonstrates the efficacy of the updated procedure and propose its suitability as a stressor task element that can be used as a potential replacement constituent or complementary element for the Trier Social Stress Task (TSST). The findings indicate that this task can be considered where participants are required to have multiple interactions with the TSST, and where external habituation effects, such as rehearsal effects - for example those associated with public speaking - can expose participants to habituation in this stressor paradigm. This original study has been published in The Journal of Psychophysiology and is outlined in detail in chapter four

Chapter five contains, study three, "Examining the use of an online adaptation of the Day Reconstruction Method", a study examining an adaptation of the Day Reconstruction Method

(DRM) originally proposed by Kahneman (Kahneman, Krueger, Schkade, et al., 2004). The primary motivation for the inclusion of this study was to examine methods that reduce researcher and participant burden. The methodology presented here demonstrates novel means of assessing participant adherence to daily diary measures. It details observed patterns of participant adherence to their assigned times for data input, with implications for studies using DRM instrumentation, including the ability to assess adherence. In addition, it details a reduction of researcher workload that integrates the use of traditional pen and paper (if required to maintain the diary-type reconstructive experience) and removes the need for participants to return to the laboratory for a final session. The study also measures affective patterns and reports participant data for the four-day collection period. The study discusses how the ready availability of these data contribute to the possibility of integration with momentary sampling and other Ecological Momentary Assessment (EMA) platforms. Finally, it is suggested that the integration of retrospective and momentary sampling is desirable, both in terms of the methodological benefits, the quality of data collected, and also how this better reflects the concept of measurement surrounding the remembered and experiencing self, and how this dichotomy could be resolved in order to better capture human experience. This study is fully detailed in chapter five

Chapter six presents study four, “From the lab to the living room: Measuring heart rate variability in ecologically valid contexts”, and details how the work in HRV measurement in ambulatory contexts, laboratory measurement, and experience sampling are combined in the major study for this PhD. Study four extends the measurement of participants’ HRV patterns from laboratory contexts, and then utilising the previously outlined measures of HR and experience sampling, examines how they relate to ambulatory patterns across four full days of 24 hour a day measurement for each participant. This study proposes a number of hypotheses:

firstly, that participants will show differences between baseline and stressor conditions inside the laboratory; secondly that baseline and matched ambulatory measurement periods will not differ; and finally, that reactivity patterns inside the laboratory will differ from the participants' ambulatory baselines. Multilevel is modelling is used to demonstrate group and fixed effects and how participants' patterns of HRV change as they move across conditions.

In order to contextualise the individual studies, framing chapters are presented in chapter two and gives a broader introduction to the modes of physiological measurement primarily examining heart rate. To begin with descriptions of the physiological basis of heart rate measurement, and in particular heart rate variability, along with its respective indices are presented. In addition, a number of relevant theoretical frameworks that have emerged using HRV measures are discussed, including, Polyvagal Theory (PVT), and Neurovisceral Integration (NVI). Also described are the applications of these measures in laboratory testing including a description of laboratory stress testing. Finally, in the introductory chapter describes existing research in experience sampling methods, including momentary sampling and diary type measures, providing the context for presented studies. To complete the thesis a broader discussion is presented in chapter seven.

1.2 Heart rate measurement

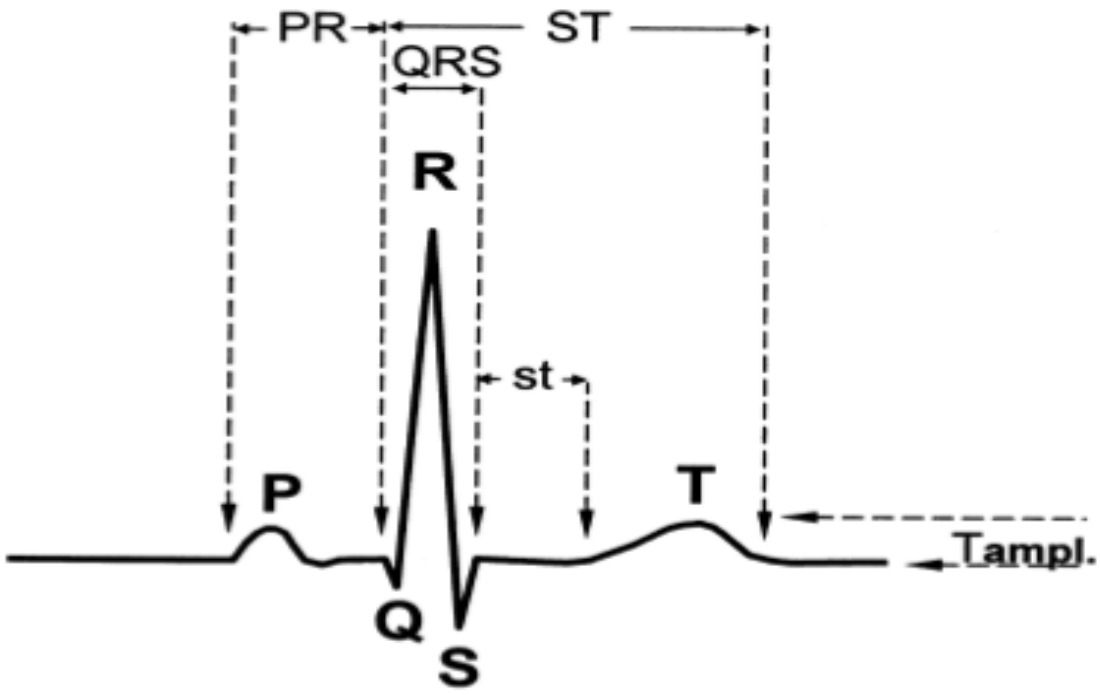
1.2.1 Heart rate measurement general introduction

The Electrocardiogram is a measurement derived from the electrical output of the heart which is produced as a result of the distinct physiology of the cardiovascular system (Berntson & Quigly, 2007). The cardiovascular system is the structure through which the body distributes oxygenated blood throughout the tissues of the body and oxygen is acquired via pulmonary circulation and distributed via arterial circulation. A secondary, venal network collects waste products and deoxygenated blood that is returned to the heart and subsequently pulmonary circulation where carbon dioxide and other waste products are removed, and the cycle begins again. The heart is comprised of specialist cardiac muscle fibres of which there are three types; atrial, ventricular, and conductive fibres, that differ from skeletal muscle and are found nowhere else in the body and form the structure of the heart. The distribution of the conductive fibres across the chambered structure of heart determines the way its musculature is mobilised to successfully coordinate the complex set of actions that comprise the heartbeat and subsequent pumping action known as the cardiac cycle. In addition to their distribution, the manner in which depolarisation of the nerve fibres occurs plays a vital role in the rhythmic nature of the cycle. Unlike skeletal muscle, depolarisation of cardiac muscle has a characteristic plateau, owing to the mobilisation of combined fast sodium and slow calcium channels (or a short period 0.2-0.3 seconds) that results in a longer contraction of the fibres. These conductive patterns result in a more efficient pumping pattern, allowing for timed release of blood from the atria, and give the heart beat its characteristic rhythmic pattern and subsequent electrical imprint.

The electrical imprint of the heart is captured at the sinoatrial node using electrocardiography (ECG) that can be observed as the characteristic PQRS(T) trace. Measurement is typically

carried out using multi-lead stationary methods. The PQRS(T) complex represents the two main epochs of the cardiac cycle - diastole and systole. The start of periodic cycle is typically marked at the beginning of arterial systole, with the depolarisation of the arterial muscle represented by the P wave. Shortly after arterial depolarisation, arterial contraction occurs followed by ventricular contraction. These patterns are characterised by the QRS sequence. Once depolarisation occurs, the diastolic phase begins, and once complete the sequence begins again.

Figure 1.1: *The PQRST trace.*



To partition the timing of this cycle the distance (and subsequently time) between successive R peaks is used as the standard measure. The R-R interval is used owing to its larger and more easily defined spike compared to other ECG trace peaks (Berntson, Quigly, & Lozano, 2007).

The R-R interval forms the basis for measurement of heart rate period (in msec) which are used to derive measures such as Beats Per Minute (BPM), and more detailed Heart Rate Variability (HRV) indices. It is the patterns and variation of these electrical imprints that is of interest to psychophysiologists because in addition to being influenced by myriad metabolic and systemic demands, and being indicative of physical health, they are also subject to change, via neural structures, that both mediate and are mediated by psychosocial variables (Porges, 2001). Observations of patterns of PQRS(T) traces have been widely examined in psychophysiological science primarily heart rate reactivity, which examines timed intervals of the r-r interval (BPM), and heart rate variability (HRV) which represents a range of time domain and frequency measures.

Patterns of reactivity have been described by the reactivity hypothesis, which posits that heightened cardiovascular reactivity, marked by increases in heart rate can be observed in response to stress, and produces changes to the structure and functioning of the heart over time. Sustained hypertensions (Krantz & Manuck, 1984), if prolonged or exaggerated can promote the development of cardiovascular disease, and are linked to the pathophysiology of a wide range of disease, including atherosclerosis, myocardial infarction and coronary heart disease, and increased mortality (Allen et al., 1997; Everson et al.,1997). The reactivity hypothesis has been extended to include the study of a wide range of psychosocial risk factors, and the impact of psychosocial variables on reactivity has been widely investigated. For example research has demonstrated how factors such as personality disorders (Eikeseth et al., 2020), psychopathy and aggression (Lorber, 2004), and antisocial behaviour (Fagan et al., 2017) are associated with patterns of cardiovascular reactivity that give insight into the mechanisms by which stress influences disease aetiology and progression.

In addition to the magnitude of reactivity, the hypothesis also considers the role of prolonged reactivity to stress, namely the duration of a stress response before cardiovascular function returns to baseline level. This phenomenon is known as recovery and has been shown to predict long term changes in blood pressure and pathology (Stewart & France 2001). A range of variables have been found to influence recovery, e.g., Larsen and Christenfeld (2011) demonstrated that an assortment of cognitive (e.g., ruminating) and behaviours (e.g. anger expression) can influence recovery. The consideration of prolonged exposure to stress and patterns of recovery generates questions related to the primary means of study employed in examining reactivity, namely laboratory-based study.

Although initially increased reactivity to stressors was primarily linked to negative outcomes, it is now known that the impact of both physical and psychological risk factors on cardiovascular health is more complex. Phillips (2013) suggests that some risk factors associated with CVD (e.g. depression and obesity) do not appear to increase risk through increased cardiovascular reactivity; in fact, lower or blunted cardiovascular reactivity can prospectively predict lower self-reported health, increased depressive symptoms, and increased risk of obesity (Phillips, 2011). Blunted CVR and neuroendocrine response have adverse consequences for health and have a number of distinct risk factors that may be behaviourally mediated. In addition, biological and cognitive factors, such as increased basal levels of arousal, that reduce the range of cardiovascular and neuroendocrine responses, and motivational dysregulation caused by prolonged exposure to stress may be contributory factors. It is likely that a specific range of reactivity, with high and low reactivity indicating maladaptive deviations from the norm, form the basis of optimal psychophysiological function. Therefore, studies that model stressors that elicit a range of responses are desirable.

The current PhD presents two studies that include reactivity as a measure. In study two, data is presented relevant to laboratory reactivity in relation to a laboratory stressor paradigm. Measures of reactivity (BPM) are included in the studies presented here in laboratory contexts and in study four in ambulatory contexts. These studies are extended to include HRV measures and ambulatory contexts examining the question of whether patterns of reactivity observed in the laboratory are replicated in ambulatory contexts. In addition to reactivity, further measures of HRV are included in the current work, as HRV offers further insight into autonomic function. The majority of studies examining reactivity have taken place in laboratory contexts following a Resting, Reactivity, Recovery (RRR) study design. While advantageous in offering control over the myriad variables that influence heart rate, for example movement, ecological validity is questionable in laboratory paradigms. Ecological validity refers both to the quality of data, collected both in the laboratory or ambulatory contexts, and whether these data are generalisable to everyday life (Shiffman, Stone, & Hufford, 2008). For example, does the observed response to public speaking observed in laboratory tasks, including baseline and the range of reactivity observed in the laboratory represent a normal range for the individual in real life. This is uncertain, and study that seeks to examine these patterns outside of the laboratory is desirable.

1.2.2 Overview of Heart Rate Variability Measures

Further to the development of reactivity measures, heart rate variability (hereafter HRV) has been proposed as an index of cardiovascular and autonomic function, and HRV measures have been widely used in psychophysiological study (Laborde et al., 2017; Quintana & Heathers, 2014). Proponents of HRV type studies suggest that the more detailed measures better describe the physiological mechanisms and underlying interdependent regulatory processes that contribute to psychosocial functioning (Porges, 1994). Although ECG data are relatively easy

to capture with equipment capable of recording the full R-R beat to beat interval, HRV data modelling requires more sophisticated analytical methods to parse the data into the respective time and frequency derivations, where R-R peak interval is used to calculate the variability of timing of the heartbeat.

Time domain measures and spectral analysis of frequency outputs have linked vagal tone and sympathetic/parasympathetic balance to affect (Kreibig, 2010; Porges, 2001) and social functioning (Porges, 2001; Quintana et al., 2012). In contrast to BPM, time-domain measures of variability use the Inter-Beat Interval (IBI) to describe changes in successive R –R peak intervals. Derivations including Standard Deviation of R_R Intervals (SDNN) and Root Mean Square of Successive Differences (RMSSD) reflect the beat to beat variance in IBI for cardiovascular activity and has been associated with frequency measures (Shaffer & Ginsberg, 2017). Frequency domain measures are based on the assumption that each range (detailed in table 1.2, pg. 30) is indicative of the 'Power' or signal energy for a given band and are described as following; Ultra Low Frequency (ULF), Very Low Frequency (VLF), Low Frequency (LF) and High Frequency (HF). ULF has been associated with circadian rhythm and metabolic demands (Shaffer et al., 2014). VLF has been linked to physical movement (Bernardi et al., 1996) and parasympathetic activation (Taylor et al., 1998). The LF band has been associated with baroreceptor activity, and subsequently blood pressure (Shaffer & Ginsberg, 2017) and proposed as an index of sympathetic outflow (Reyes et al., 2013) although inclusion of LF, and subsequently LF/HF ratio, is questionable and disagreement exists regarding the use of either measure as an index of sympathetic balance because considerable disagreement across studies exists (Billman, 2011, 2013; Heathers & Goodwin, 2017). HF components of HRV have been reliably associated with vagal tone (Porges, 2007), and correlated with stress, permitting its use as a measure to reflect parasympathetic activity (Laborde et al., 2017). Finally, low frequency

/ high frequency ratio has been proposed as an index of autonomic outflow (ChuDuc, NguyenPhan, & NguyenViet, 2013; Quintana & Heathers, 2014). Although as mentioned because the inclusion of LF bands is questionable, LF/HF ratio carries similar concerns. For the purpose of studies presented here, time domain measures of BPM, IBI, RMSSD and frequency components of HF are included, with the exclusion of LF and LF/HF ratio.

Table 1.1: *Time domain measures of HRV- (Shaffer & Ginsberg, 2017)*

Parameter	Unit	Description
SDNN	ms	Standard deviation of NN intervals
SDRR	ms	Standard deviation of RR intervals SDANN
SDANN	ms	Standard deviation of the average NN intervals for each 5 min segment of a 24 h HRV recording
SDNN index	ms	Mean of the standard deviations of all the NN intervals for each 5 min segment of a 24 h HRV recording
pNN50	%	Percentage of successive RR intervals that differ by more than 50 ms
HR Max- HR Min	bpm	Average difference between the highest and lowest heart rates during each respiratory cycle
RMSSD	ms	Root mean square of successive RR interval differences
HRV Triangular index		Integral of the density of the RR interval histogram divided by its height
TINN	ms	Baseline width of the RR interval histogram

Table 1.2: *Frequency domain measures of HRV- (Shaffer & Ginsberg, 2017)*

Unit	Parameter	Description
ULF Power	ms^2	Absolute power of the ultra-low-frequency band (≤0.003 Hz)
VLF Power	ms^2	Absolute power of the very-low-frequency band (0.0033–0.04 Hz)
LF Peak	Hz	Peak frequency of the low-frequency band (0.04–0.15 Hz)
LF Power	ms^2	Absolute power of the low-frequency band (0.04–0.15 Hz)
LF Power	nu	Relative power of the low-frequency band (0.04–0.15 Hz) in normal units
LF Power %	%	Relative power of the low-frequency band (0.04–0.15 Hz)
HF Peak	Hz	Peak frequency of the high-frequency band (0.15–0.4 Hz)
HF Power	ms^2	Absolute power of the high-frequency band (0.15–0.4 Hz)
HF Power	Nu	Relative power of the high-frequency band (0.15–0.4 Hz) in normal units
HF Power %	%	Relative power of the high-frequency band (0.15–0.4 Hz
LF/HF	%	Ratio of LF-to-HF power

1.2.3 Epoch duration and selection for HRV measurement

ECG data used to derive HRV measures can be captured across different timeframes using a number of methodological approaches, including long term (24 hr +), ultra-short term (1minute), and short-term (less than 24 hrs comprised of 5-minute epochs). Shaffer (2017) provides details on normalised values for matched time durations and summarises the range of metrics and norms derived from each context (Shaffer & Ginsberg, 2017). Short term measures, are described as those of between five minutes and twenty-four hours, typically parsed into standardised epochs of five-minute duration (Camm et al., 1996; Laborde et al., 2017). Nunan (2010) reports values for short term HRV, including; Total power (TP) SDNN, RMSSD,

pNN50%, VLF LF, HF and LF/HF ratio (Nunan et al., 2010) and provides recommendations for study designs to align with short-term designs. Short term recordings are included in the presented studies here for matched comparison of laboratory stressors and ambulatory contexts and the norms and parameters outlined by Shafer and Nunan were used to inform the data analysis.

Shorter duration recordings (Ultra short term >1min) carry advantages such as convenience and time duration while sampling and can be utilised in HRV studies (Esco & Flatt, 2014) with shorter duration sufficient to generate a reliable measurement. However the reliability of UST duration recordings are questionable as they are more susceptible to methodological error (Shaffer & Ginsberg, 2017), and although concurrent validity with other epoch durations has been demonstrated they incur a high workload at data analysis. In order to account for artefacts, non-contiguous segments can be combined over a study design (Shaffer et al., 2016), and longer (5min) durations are desirable to standardise measurement and to allow accurate calculation across HRV parameters (Laborde et al., 2017). Long term measures, or those characterised as longer than 24 hrs can be constructed in a number of ways: firstly global, that is HRV calculated over the entire 24 hr period, and secondly 24 hr times frames parsed into five-minute segments. Laborde (2017) recommends that 24hr recordings be constructed based on contiguous measurements of five-minute epochs and utilised in the long-term ambulatory study. The data presented here, in the two studies concerned with HRV measurement follow these recommendations. Firstly, in the laboratory-based paradigm in study two (Jump & Dockray, 2020) the five-minute epoch structure is demonstrated within a Resting, Reactivity, Recovery (RRR) design in a laboratory context. Further to the laboratory study, the five-minute segmented epoch structure is employed over the four-day measurement period in study four. This approach gives opportunity to compare measures across contexts, and this study aimed to

demonstrate its application using a selected number HRV parameters. Comparative measurement in this context raises a number of methodological challenges and for each study there are considerations related to inclusion and exclusion of measures. The HRV measurements that are included are detailed in the presented studies respectively

1.2.4 Tonic and phasic HRV

Tonic HRV refers to the measurement of resting or baseline HRV. Tonic measurement may be carried out in isolation or in conjunction with reactivity and recovery periods and has been associated with health (Kemp & Quintana, 2013), affective regulation (Berna et al., 2014; Duarte & Pinto-Gouveia, 2017), and psychosocial variables (Porges, 2001) In general, higher tonic HRV levels, indicating a wider range of variability, is associated with better health outcomes and has been implicated as marker of the dynamic capacity of the individuals' physiology and subsequently their ability to react appropriately to environmental stimulus. Individual differences in tonic HRV have been examined primarily using laboratory-based measurement (Laborde, 2017). Typically, self-report measures, are used to categorise individual differences in either a state or trait type measures and these will be associated with resting HRV.

Phasic HRV refers to the change in HRV in response to challenge and has been termed vagal withdrawal (Park et al., 2013), for example to a psychosocial stressor such as a public speaking task. While in general higher tonic HRV is associated with positive health outcomes the adaptive relationship of phasic HRV to health is more complex than tonic HRV and the context and magnitude of the phasic response is important - for example, if a high magnitude response takes place in reaction to a dangerous or physically threatening situation and vagal withdrawal occurs, then the resultant increase in heartrate and energy levels provided can be viewed as

adaptive (Porges, 2007). However, if high magnitude phasic responses occur inappropriately, for example, relation to a social situation like public speaking, and the individual is dysregulated and unable to appropriately respond, then it can be viewed as maladaptive (Geisler et al., 2013). The measurement of phasic HRV should therefore be contextually informed and efforts made to account for confounds. The dominant theoretical frameworks, namely polyvagal theory (Porges, 2001) and Neurovisceral integration (Thayer & Lane, 2000) (further detailed in section 1.3) provide a rationale for measuring tonic and phasic HRV, and in particular their relationship to social context.

The studies presented here examine the use of tonic and phasic HRV in both the laboratory and ecologically valid settings to interpret their association with social context and to describe the challenges associated with measurement. The focus of the thesis presented here is the presentation of study four examining the application of HRV in ambulatory contexts. Although the study of HR in laboratory contexts is rich and has been linked to a wide range of outcomes, ambulatory study in contrast is relativity sparse. The reasons for this are varied, but mainly owing to the complexity of ambulatory measurement and the myriad confounds that exist once participants are out of the controlled environs of the laboratory. Although difficult, these measurements are not impossible; Laborde (2017) suggests the controls required to capture HRV data in ambulatory contexts, including the use of accelerometer data to account for movement and signal quality (Smets et al., 2019).

1.3 Theoretical frameworks

1.3.1 Polyvagal Theory

"When we view living organisms as a collection of dynamic, adaptive, interactive, and interdependent physiological systems, it is no longer appropriate to treat the autonomic nervous system as functionally distinct from the central nervous system. We start to recognize that peripheral organs do not "float in a visceral sea", Rather, they are anchored to central structures by means of efferent pathways and are continuously signalling central regulatory structures along their abundant afferent pathways" (Porges, 1994)

Polyvagal theory (Porges, 1994) is built on a substantial body of work prior to its introduction in 1994 and since then a large number of studies have examined its various claims. Porges' early studies, prior to the proposal of Polyvagal theory, took a number of major strands including (but not limited to), firstly, a series of experiments in which heart rate variability was used as a measure to examine existing work, (Lacey, 1967) for example, on how respiration and heart rate interact with attention, in particular how attention is mediated by autonomic activity (Porges & Raskin, 1969), the influence of respiration on heart rate (Cheung & Porges, 1977; Fast & Porges 1970), and reaction times (Porges, 1973). The second major strand was concerned with the developmental characteristics of HRV in children and infants, including; Individual differences in cardiac activity (Fracasso et al., 1994), sensory input such as auditory stimuli (Forbes & Porges, 1973), visual input (Porges, Stamps, & Walter, 1974), relationships between heart rate and attention (Porges, 1974; Porges et al., 1973; Suess et al., 1994), hyperactivity (Porges & Smith, 1980) childhood psychopathology (Greenspan & Porges, 1984) and autism (Porges, 1976). The third strand in Porges' work examined physiological measures such as blood pressure, heart rate reactivity and heart rate variability and their functional relationship to internal states (Hatch, Klatt, Porges, Schroeder,Jasheway, & Supik, 1986;

Porges, 1985; Porges, 1992), general nervous system state and the regulation of emotion (Porges, 1988). The fourth major strand to emerge from Porges' work, which was key to the development of Polyvagal theory, outlined the procedures that developed during the course of his work, including; designs for infant research (Porges, 1979) developments in statistical applications to HRV analysis (Porges et al., 1980; Porges & Bohrer, 1990), descriptions of the considerations of respiratory sinus arrhythmia and its influence on HRV (Byrne & Porges, 1993; Porges, 1986; Porges, 1986) and research methods for heart rate measurement (Porges & Byrne, 1992). It was this body of work that culminated in the emergence of Polyvagal Theory.

Polyvagal Theory was first proposed in 1994 in "*Orienting in a defensive world: Mammalian modifications of our evolutionary heritage. A polyvagal theory*" (Porges, 1994), and provides an explanation for how autonomic function relates to behaviour with a focus on the neural regulation of the heart via afferent and efferent pathways of the vagus nerve (Porges, 2007). The model was an attempt by Porges to move away from stimulus response (S-R) models of heart rate measurement in psychophysiology and to locate HRV research within a neurophysiological framework. Previous models had shied away from describing neural regulation of the heart and even viewed HRV signals that were detected during experimentation as a type of measurement error. Explanatory models at the time such as those posited by Lacey, focused on mediation of heart rate via baroreceptors, (Lacey, 1967), and Obrist who proposed the changes in heart rate were a solely a function of metabolic demands (Obrist, 1981). These models represented the prevailing thought at the time which seemed to ignore any top-down neural input. Porges summarised his objection to these models by saying,

"These theoretical positions, if taken to an extreme, would assume that alterations in heart rate would be entirely under the control of experimental or environmental demands. This, of course,

is in direct contradiction to our current knowledge of the dynamic relationship between the heart and the central nervous system and the influence of this interaction on the production of HRV" (Porges, 2007, p5).

Polyvagal theory began the process of "anchoring" HR responses to the environment within the central nervous system, focusing on the morphology of the vagus nerve and subsequently the related functional and behavioural outcomes. In addition, it provided an initial model of how neural structures related to psychological phenomena, such as social functioning. The theory outlines how heart rate is altered by psychological responses to environment, such as stressors or novelty and views HRV as the "Superimposed sum of several rhythmic oscillations and slow trends that co-vary with metabolic demands" (Porges, 2007, p7).

Central to polyvagal theory is a description of the anatomy of the vagus nerve. The vagus nerve is the tenth cranial nerve and has the longest, most complex distribution of all the cranial nerves and subsequently has a diverse functionality. Although often conceptualised as a single nerve the vagus has several branches that include afferent and efferent pathways. The nerve is comprised of five different types of nerve fibre including, general somatic afferent, general visceral afferent, special visceral afferent, general visceral efferent, and special visceral efferent. The vagus originates in a lateralised manor in the medulla and distributes to enervate structures in the thorax, neck and abdomen, and also including the heart. (Câmara & Griessenauer, 2015). Of particular interest in polyvagal theory are the two primary source nuclei in the medulla - the nucleus ambiguus, and the dorsal motor nucleus. Polyvagal theory seeks to describe how the anatomy of the vagus nerve relates to behaviour, in fact it was this that defined Porges' work on HRV.

One of the key features of Polyvagal theory is its limitation of enquiry to this single structure while acknowledging the myriad afferent and efferent influences on HRV measurement. The theory emphasises how the anatomical differences in these pathways support adaptive behavioural outcomes and how these anatomical differences have been layered in a hierarchical manner that have been mediated by evolutionary processes. Porges summarises these differences in three phylogenetic stages as follows; Firstly, the unmyelinated vagus mediates immobilisation via the dorsal motor nucleus. Secondly the sympathetic adrenal system activates avoidance behaviour via the spinal cord. Thirdly the myelinated vagus is responsible for social communication, arousal inhibition and deactivation via the nucleus ambiguus. Which system is activated and prioritised is determined by the specific environmental challenge at hand, for example; if no immediate physical danger is present, that requires intervention, such as freezing behaviour or immediate withdrawal (or other action) by the unmyelinated vagus or the adrenal system (Systems 1 & 2) then heart rate is mediated via the myelinated vagus (system 3). This description of these systems by Porges of adaptive reaction to challenge was a key feature of Polyvagal theory.

In addition to patterns of activation across unmyelinated nerves and the endocrine system the vagus nerve provides the mechanism by which inhibition (or deactivation) of arousal occurs. Porges refers to this mechanism as the vagal brake. The ‘vagal brake’ is a description of how the myelinated vagus acts to disengage and override the arousal patterns in the unmyelinated vagus and the endocrine system, and also how this deactivation is mediated. The myelinated vagus, specifically because it is myelinated, and therefore physically separated from the surrounding nerve structures, can act independently. It follows that the physiology of the myelinated vagus reflects the ability of the vagal brake to act as one of the ways arousal is regulated, and this ability is reflective of the individuals’ ability to regulate their visceral state.

The vagal brake is mediated via enervation of the myelinated vagus in the nucleus ambiguous and the origin of the nerve has an isolated and dedicated neurology that enables differential activation, and these neural structures are also subject to top-down regulation from higher cortical regions. It is this distinct physiology that not only mediates the visceral state and subsequently the ability of the individual to navigate complex social environments but also explains individual differences in regulation of the visceral state. As with other systems, for example hypothalamic pituitary adrenal (HPA) activation, the patterns of vagal activation, and the ability of the vagal brake to regulate HR, represent a life-long activation profile for the individual that can be indexed via HRV measures and is correlated with a host of psychosocial variables. It is this distinction that has given rise to the body of work that has emerged using Polyvagal theory as a scaffold for enquiry. By delineating physiological structure and specifying HRV parameters incurred via environmental challenge, and subsequently its relationship to psychological state, it has allowed nuanced study of the relationships to these challenges. This is a key area that has defined Porges' work, namely the examination of how individual differences in functionality of the vagus determines the regulation of social behaviour. Porges posited that "the phylogenetic origin of the behaviours associated with the Social Engagement System is intertwined with the phylogeny of the autonomic nervous system" (Porges, 2007, p10). In other words, the evolution of the social context in which humans emerged is inextricably linked to the neural structures that developed to meet the demands that were placed upon them.

These structures include the enervation from the face and head that are linked to the upper cortices and the myelinated vagus via the nucleus ambiguus, and by implication represent some of the top-down pathways by which physiological reactivity is mediated in social interaction.

This mediation is bi-directional and because these structures have a somatosensory component they allow ‘feed-back’ information to be sent to the sense organs which in turn is indicative of the internal emotional state. This ability to effectively determine one’s own internal visceral state informs how self-regulation occurs and allows the individual to align their internal state with the environmental context, for example to accurately assess risk and then appropriately respond. This ability is key to determining appropriate levels of physiological reactivity to context via vagal activity, and consequently ensuring appropriate social behaviour in a given context. In Polyvagal theory it is this ability to regulate physiology appropriately that defines how organisms will navigate complex social environments. In effect, Porges described the mechanism by which (parts of) the internal visceral state of the organism interacts both with their environment and the individuals that inhabit it. Porges also offers explanations for maladaptive responses in this mechanism as a possible explanation for a range of pathologies including, anxiety, depression and features of social functioning in autism spectrum disorders (Geisler et al., 2013; Stephen W. Porges, 2001)

Emerging from this work Porges outlines four principles of social engagement that provide the framework for the enquiry into the physiological mechanisms that mediate social engagement. Porges summarises these principles by stating “the mammalian autonomic nervous system provides the neurophysiological substrates for the emotional experiences and affective processes that are major components of social behaviour” (Porges, 2007, p9).

Table 1.3: *Principles of social engagement (Porges, 2007)*

Principal	**Description**
1	There is a phylogenetic shift in the regulation of the heart from endocrine communication, to unmyelinated nerves, and finally to myelinated nerves.
2	There is a development of opposing neural mechanisms of excitation and inhibition to provide rapid regulation of graded metabolic output.
3	A face-heart connection evolved as source nuclei of vagal pathways shifted ventrally from the older dorsal motor nucleus to the nucleus ambiguus. This resulted in an anatomical and neurophysiological linkage between the neural regulation of the heart via the myelinated vagus and the special visceral efferent pathways that regulate the striated muscles of the face and head.
4	With increased cortical development, the cortex exhibits greater control over the brainstem via direct (e.g., corticobulbar) and indirect (e.g., corticomedullar) neural pathways originating in motor cortex and terminating in the source nuclei of the myelinated motor nerves

1.3.2 The Neurovisceral integration model

The Neurovisceral Integration Model (NVI), introduced by Julian Thayer in, '*A model of neurovisceral integration in emotion regulation and dysregulation*', (Thayer & Lane, 2000) attempts to integrate and account for the networks that serve to enable " the complex mix of cognitive, affective, behavioural, and physiological concomitants of normal and pathological functioning" (Thayer & Lane, 2000, p 202). Thayer posited that understanding functional aspects, such as attention and affect, along with structural elements of the nervous system and the cardiovascular system, would enable a more thorough description of how these systems form the basis of the individual's response the environment and serve as a framework for further enquiry. Since its introduction in 2000 this model has garnered a considerable amount of attention (Thayer et al., 2012) and subsequently empirical support, via a large body of research. "There is now very compelling evidence that parasympathetic or vagal tone − often measured using the high frequency component of heart rate variability (HRV) − is associated with a variety of psychological and behavioural variables on the one hand, and a variety of health outcomes on the other" (Smith, Thayer, Khalsa, & Lane, 2017, p 275).

The NVI model posits structural components (or networks) and in particular how frontal and mid brain regions are related to top down function that exerts control over the heart via the vagus nerve (Thayer, Åhs, Fredrikson, Sollers, & Wager, 2012), These changes can be captured via HRV measurements that allow for specificity to environmental, behavioural, and psychological variables. Prior models, such as Polyvagal theory (Porges, 1994; Porges, 2007) had been limited to descriptions of the anatomy of the medulla, sensory organs and the vagus nerve. The Central Autonomic Network (hereafter CAN) (Benarroch, 1993), described functional aspects of the anterior cingulate cortex, ventromedial prefrontal cortices, the amygdala, hypothalamus, medulla and insular cortex. Thayer's work extended these

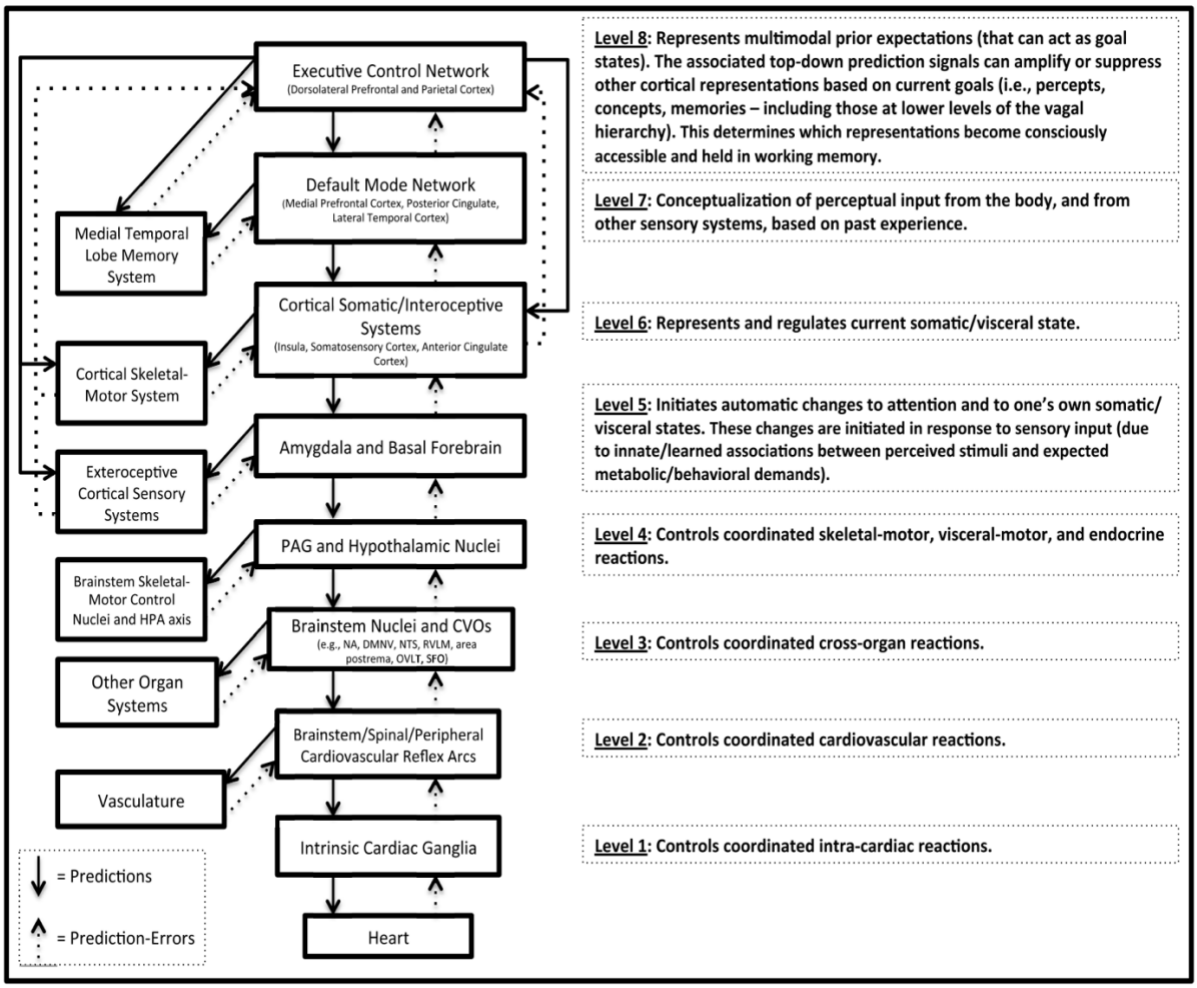
descriptions and elaborated on them, initially utilising a "dynamical systems perspective" (Thayer & Lane, 2000, p 203) to describe how the CAN is directly indexed by heart rate variability via enervation of the vagus nerve and the sinoatrial node and that this operates as a 'dynamic system' with bi-directional feedback. The suggestion being that these interactions represent executive function, social capacity, attention and affect for the individual. In 2009, in a paper updating the propositions of the model and presenting supporting research, Thayer goes further to suggest "This sparsely interconnected neural complex allows for maximal organism flexibility in accommodating rapidly changing environmental demands. When this network is either rigidly coupled or completely uncoupled, the ability to recruit and utilize appropriate neural support to meet a particular demand is hampered, and the organism is, thus, less adaptive" (Thayer et al., 2009, p143). This implies that the myriad interconnected systems are the mechanisms by which human beings successfully navigate the social environment, and crucially, that HRV can be used to index this adaptability that can serve as a window into its functional capacity.

While initially it was believed that cognitive function would relate directly to HRV, when further examined it appeared that there was specificity between HRV and the different aspects of executive function and that these relationships were not uniformly robust (Jennings, Allen, Gianaros, Thayer, & Manuck, 2015), leading to the examination of the myriad functional capacities associated with HRV. The functional capacity of these systems relates to how the individual "continuously assess[es] the environment for signs of threat and safety and to prepare the organism for appropriate action" (Thayer et al., 2012). The 'appropriate actions' were thought to primarily describe how HRV is linked to attentional and affective responses that relate to the inhibition of the distributed networks, such as those in the prefrontal cortex involved working memory. For example, inhibition of affect and directed attention are required

to successfully engage in goal directed behaviour and optimise capacity. Evidence for this relates primarily to resting HF- HRV. The role of filtering attention and working memory had been explored prior to NVI (Shimamura & Shimamura, 2000), and what the model contributed was the inclusion of HRV, as an index of these functions.

Further to the initial proposal of NVI as a dynamic systems perspective, Thayer expanded the model to describe the manner in which the networks are organised hierarchically in a multi-level structure and details how each level exerts vagal control based on context (Smith, Thayer, Khalsa, & Lane, 2017). This multilevel structure (Figure 1.2) represents how vagal control is influenced by the generation or integration of signals from related systems and how complexity increases at higher levels of the model. For example, at Level 1: Intra cardiac control, chlorogenic and adrenic neurons exert opposing influence on heart rate and represent automated homeostatic processes. At Level 5: visceral, somatic, and cognitive/attentional responses are mediated via the amygdala, and the model describes how the connections to both cortical regions and to lower brain structures, enable responses to external environmental stimuli. This allows a description of the mechanisms that generate affective response, for example to threat or uncertainty. The advantage of the hierarchical model is that it gives preference to how activation and inhibition of HR via the vagal brake is mediated in a stepwise manner, that allows for specificity to neural structures. In addition, it begins to describe the myriad points of convergence of neural input each contributes to homeostasis, and how each level exerts control over heart rate as neurology is recruited.

Figure 1.2: *Multilevel NVI structure* (Smith et al., 2017)



In addition to describing the gross differences between visceral, and higher cortical regions, such as level 3, which involves brainstem nuclei and describes how organ specific visceral input contributes to homeostasis, and level 6 that regulates homeostasis based on perceptual inputs, the model provides a granularity that allows specificity to behaviour, affective regulation, such as how experience and sensory inputs are integrated to form individual differences in responses to the environment. For example, Level 7 describes how perceptual and sensory inputs are integrated with past experience and conceptual recognition, and linked to the medial pre frontal cortex to HRV (Thayer, Åhs, Fredrikson, Sollers, & Wager, 2012). Thayer suggests that at this level the interpretation of HRV signals can be associated with the

cortical regions responsible for integrating visceral and bodily functions. This is distinct from level 8 that integrates top-down inputs or an "executive control network". This network represents working memory and allows information to be consciously held and integrated via a "top-down amplification" of the lower-level processes. For example, while level 7 is responsible for the integration of sensory environmental challenge, level 8 implicates neurology that provides context and guides goal directed behaviour. However, it should be noted that while vagal control is linked to mental process at these higher levels there are better correlates with lower order regions. Thayer suggests this indicates that although the model is hierarchical, vagal control is still mediated via the lower levels first. For example, while frontal parietal regions have been linked to HRV, this only represents "indirect control of emotion".

In a reflection of the earlier dynamical systems perspective the hierarchical model requires homeostasis or balance at each level before it can describe more complex function. To take the previous example, if at level 7 there is a pathology that interrupts the individual's ability to successfully integrate past experience and recognition then this will relate to HRV measures. In this context an individual will have difficulty in recruiting level 8 process to mediate and contextualise in a top-down regulation. More importantly, and this relates to the work presented here, Thayer suggests that these mediation processes between levels are contextual and environmentally driven. For example, in order to appraise and mediate information, recruitment of memory that contextualises the information is required at specific levels of the model. If the individual is in a novel environment, for example a psychophysiology laboratory, then this could have implications for the ecological validity of the measurements.

1.3.4 Summary

Following the task force on HRV research in 1996 (Camm et al., 1996) that sought to summarise and guide how HRV research was conducted a number of dominant research threads emerged. Firstly, biomedical research focused on the technical requirements of HRV measurement, which contributed to detailing how the various time and frequency domains were derived and in particular how confounds were managed. Secondly, research within the medical literature, concerned with disease aetiology and clinical applications (Sassi et al., 2015). Finally, research within psychophysiology that aimed to delineate how HRV was associated with psychosocial variables, such as affective regulation (Balzarotti et al., 2017), psychopathologies (Beauchaine et al., 2019), and health (Chambers & Allen, 2007). Thayer and Porges' theories emerged as the dominant frameworks in this strand, and Polyvagal theory and Neurovisceral Integration have been the focus of extensive research examining HRV and the social context of humans. Porges first provided an evolutionary explanation of how individuals navigate complex social interactions and related this to the physiology of the vagus, and subsequently the derived HRV indices. Thayer elaborated on this work, initially with descriptions of the myriad ways HRV is affected via top down and higher order neurological structures, along with the central autonomic networks, and began the descriptions of how these systems integrate with each other, through a dynamic systems perspective and later in a hierarchical model.

In order to inform these theories an extensive literature has emerged and the underlying principle, that HRV can be used as a marker of top down and affective regulation has been supported (Holzman & Bridgett 2017). However it should be noted that the studies examining HRV in relation to these theories have primarily been carried out in laboratory contexts with a focus on measuring resting/tonic and phasic/reactive HRV and associated psychosocial

variables (Laborde et al., 2017). This is owing primarily to the requirements to control for confounding variables and to allow the manipulation of isolated contexts, for example, public speaking tasks. Study Two (Jump & Dockray, 2020) presented here, addresses a stressor paradigm that examines resting, reactive and recovery in a laboratory context. This study locates the work presented here in a laboratory context and demonstrates the use of a novel social stimulus that adheres to measurement standards for HRV analysis. It also examines the social context of this type of measurement and reports tonic and phasic HRV in relation to an anticipatory stressor and has relevance for the specificity to social context. The study also provides the basis for comparison of HRV pattern in the subsequent ambulatory study presented here.

Given the importance of social context, ambulatory measurement of HRV and its psychosocial correlates is of increasing interest for a number of reasons. Firstly, it can give a better prediction of how situationally located and prolonged exposure to stressors result in HRV measurements which are linked to health (Smith, Deits-Lebehn, Williams, Baucom, & Uchino, 2020). Secondly, behavioural mediation plays an important role in people's exposure to stressors (McEwen, 2001), for example, public speaking tasks that are used in laboratory testing, may be completely avoided usually. Therefore, how people navigate their own everyday environments may be more reflective of how they interact with or avoid stressors. Thirdly, as technological advances allow more people to measure and interpret their own physiological data, a baseline of research is required to contextualise and authenticate the various claims made related to HRV in these contexts. Ambulatory research using HRV in association with psychosocial measures is therefore desirable, however, measurement outside of the laboratory carries myriad methodological concerns. Brown provides a review of the approaches to measuring HRV in everyday life (Brown et al., 2017) and others have proposed

statistical modelling techniques to account for confounds such as movement, for example using accelerometer data, and propose that it is possible to discriminate between metabolic and non-metabolic HRV in ambulatory contexts (Brown et al., 2017) and to carry out situationally relevant HRV measurement (Laborde et al., 2017). Study Four presented here details the relationships between ambulatory and laboratory measures and attempts to define a limited means for assessing both tonic and phasic HRV in these contexts.

In relation to the theoretical frameworks presented here this type of study is important to consider as this research extends beyond the laboratory to ambulatory contexts. Although the dominant research paradigms to date have contributed widely to our understanding of HRV in laboratory contexts, there is a paucity of research examining ambulatory and laboratory research together (Smets et al., 2019). Moreover, as both Porges' and Thayers' work is heavily focused on the social context of humans, for example, Thayer specifies how neurology and function differentiate based on past experience. This raises the question, if participants are exposed to novel social contexts, such as a laboratory, does this carry implications for the relevance of the data. If measurement techniques seek to give insight into participants' HRV profiles based on experiences which have documented effects on individuals such as white coat effects (Leventhal et al., 2007). This raises the questions related to the central tenet of HRV measurement, namely how HRV measures relate to experience and the individual's psychological and behavioural propensities. Therefore, research that links HRV research in laboratory to ambulatory contexts is necessary. In the following sections we briefly discuss the types of laboratory and ambulatory measurement employed in the later presented studies.

1.4 Laboratory testing of Heart Rate Variability (HRV)

Models of psychobiological stress reactivity rely on a foundation of reliably mapping responses to standardised stress tasks. Standardised tasks are desirable as, once properly administered to account for confounds, they allow physiological and psychosocial variables to be associated (Chida & Hamer, 2008). To delineate these responses a dedicated literature has emerged examining responses patterns and their methodological application (Allen et al., 2014). For HRV research, laboratory procedures that model tonic/resting and phasic/reactive HRV and allow tailoring of differentiated types of social contexts with matched timing are most favourable. Although numerous procedures exist, the Trier Social Stress Task (TSST) (Kirschbaum et al., 1993) has emerged as the most widely used stress test in psychophysiology (Kudielka, Hellhammer, & Kirschbaum, 2007). The TSST is currently the gold standard for the induction of stress in laboratory contexts. It has been employed to elucidate differential responses in various types of psychophysiological studies, for example, neuroendocrine responses via cortisol, Adrenocorticotropic Hormone (ACTH), vasopressin and Dehydroepiandrosterone (DHEA) indicative of Hypothalamic Pituitary Adrenal (HPA) axis activation, and cardiovascular activation (Allen et al., 2017). The TSST has been employed to study stress response in a wide range of demographic and psychosocial measures including, age & gender (Kudielka, Buske-Kirschbaum, Hellhammer, & Kirschbaum, 2004) and pubertal stage, (Sumter et al., 2010), and to explore the physiological mechanisms of psychopathologies, including; depression (Burke et al., 2005; Young et al., 2000) and anxiety (Gerra et al., 2000; Young et al., 2004).

Although the TSST has been widely used in these types of studies it should be noted that a number of challenges do exist including, ecological validity (Wetherell & Carter, 2014), ceiling

effects (Allen, Kennedy, Cryan, Dinan, & Clarke, 2014) , and methodological resources load. For example, using the Group TSST requires a large number of participants and confederates. Finally, although the TSST has demonstrated efficacy in CV measures in both group and individual contexts issues regarding confounds such as movement and speaking exist for HRV type studies, leading to the proposal of novel methodologies. Resulting from the need to address these challenges novel means of stress induction, such as a paradigm for inducing mental stress - the Sing a Song Stress Test (SSST) (Brouwer & Hogervorst, 2014) - have been proposed. Although there are TSST variants that allow for the restriction of these confounds they did not introduce a novel stimulus, which can be used to address habituation effects, a typical physiological phenomenon where participants demonstrate a higher magnitude response the first exposure to a stressor and a subsequent reduction or diminished reactivity to repeated exposure (Hare, Wetherell & Smith., 2013).

The SSST elaborated on a methodology proposed by Hoffman (Hofmann et al., 2006), by using the instruction to "sing" and employed it to examine reactivity in shy/non shy participants using the anticipation of the singing event to induce the stress response. In Hoffman's original methodology participants were informed that they would be required to give a speech and sing a song. Hoffman's task was primarily designed to elicit an anticipatory stress response in reaction to a future event. It also aimed to include a novel stimulus, namely a public singing task. Brouwer further developed this task and adapted the methodology in order to standardise the presentation of the stimulus, using computer-based prompts instead of those typically spoken by the researcher. The initial findings from the SSST indicated a reliable induction of a cardiovascular reactivity stress response using an anticipatory sing a song stimulus. The methodology also had the advantage of reducing confounds such as movement and speech. While the SSST does carry advantages, particularly for short term CV measurement, a number

of methodological challenges are present for researchers considering its use with HRV measures, including; matched timing of epochs of sufficient duration for HRV measures, and adherence to a resting/reactivity/recovery procedural outlay best suited to HRV analysis (Brown et al., 2018; Laborde et al., 2017; Shaffer & Ginsberg, 2017).

In order to address these challenges and ensure this type of social stimulus could be used in the current studies, a number of adjustments were made to bring the task into alignment with recommendations (Laborde et al., 2017) for experimental planning, data analysis, and data reporting for HRV type studies, including; a Resting, Reactivity & Recovery design with anticipation of singing stressor stimuli included along with established stressor protocols, matched timing of <5 minute epochs, and the use of ECG capable devices. In addition, the use of this novel anticipatory stressor had not been examined using more granulated measures of HRV that offer the previously mentioned advantages in interpretation of autonomic function (Quintana & Heathers, 2014). Study two presented here (Jump & Dockray, 2020) details the viability of the anticipatory singing task, both as a complementary or replacement element for the standard TSST and for HRV measurement. In order to adhere to standardized methods for measurement of cardiovascular activity, including a Resting/Reactivity/Recovery (RRR) structure (Laborde, Mosley, & Thayer, 2017; Shaffer & Ginsberg, 2017) the methodology is amended as detailed in study two and suggests that the method not only allows for delineation between tonic and phasic HRV between conditions but also stressor tasks with a differing psychosocial and cognitive demands and this is detailed in study two. The inclusion of this study also forms the comparative basis for the later studies presented here that capture ambulatory contexts.

1.5 Ambulatory measurement

Although laboratory based stressors are extensively employed in physiological research measurement of participants' responses in this context can carry certain disadvantages (Chida & Hamer, 2008). These include a narrow representation of stressor tasks (Wetherell & Carter, 2014), habituation to repeated exposure to stressors (Arvidson, Sjörs, & Jonsdottir, 2017; Schommer, Hellhammer, & Kirschbaum, 2003). As mentioned in previous sections, the dominant theoretical frameworks suggest that social context is important when considering both tonic and phasic HRV. With laboratory stressor tasks this is potentially further confounded when considering the possibility that participants would otherwise behaviourally mediate exposure to stressors, so that the laboratory tasks do not represent a typical event - for example public speaking. Finally, with technology now allowing wider access to physiological data to it is important that research extends to real world contexts. The work presented here examines comparisons between laboratory measures and ambulatory measures.

Studies that employ HRV measures in ambulatory contexts can be captured across multiple modes. Firstly, taking non-interruptive long-term measures examining overall patterns in HRV, where global measures are correlated with psychosocial indicators. Secondly, specified short term measures, often linked to specific events are also used. As mentioned previously here HRV measurement can be captured across different timeframes including long term (24 hr recording), ultra-short term (UST >1min) are also possible (Esco & Flatt, 2014) but their reliability is questionable (Fred Shaffer & Ginsberg, 2017). Short-term (5-minute- 24hr) epochs (Fred Shaffer & Ginsberg, 2017) can be used to derive HRV measures (Camm et al., 1996; Laborde et al., 2017) including total power (TP) SDNN, RMSSD, pNN50%, VLF LF, HF and LF/HF ratio (Nunan et al., 2010). Finally, long term measures, longer that twenty-four hours, can be constructed over an entire 24 hr period or comprised of 24 hr times frames parsed

into five-minute segments. These measures are then adapted given the type of study design employed.

Studies employing HRV measures in ambulatory contexts apply appropriate timeframes informed by their respective designs and research questions; for example HRV patterns and job stress and burnout have been studied extensively studied using this approach (de Looff et al., 2018), with the majority of these studies examining the relationships between self-report measures and twenty four hour HRV patterns. Other approaches favour the inclusion of specified tasks in ambulatory contexts with physiological measures such as CVR and HRV. For example, Dikecligil & Mujica-Parodi (2010) conducted a comparison of a laboratory stressor, the affective picture scale, and an acute ambulatory stressor, a skydiving activity, to model a real-world stressor (Dikecligil & Mujica-Parodi, 2010). Their findings indicated good correlations between short term and longer term HRV, and that laboratory responses provided a marker for real world acute stressor reactivity. In addition, they used comparisons of the laboratory reactivity and compared it to the longer term (24 hr) recordings associated with the waiting period for the skydive. Findings for longer term HRV indicated that only values taken during the "waking" period, defined as between 7.30-23.30 were associated with each other. In addition, values are reported for LF/HF ratio, and as outlined here this carries methodological concerns. A similar type study (Hare, Wetherell & Smith., 2013) compared experienced and novice skydivers using salivary cortisol to measure HPA activation. While the groups did differ with regards to self-report affect, they found no difference between the two groups in terms of physiological reactivity to the stressor.

For shorter term ambulatory measurement of HRV others have attempted to frame and control the measurement periods, for example Gerin et al (1993) describe, with some caveats,

how reproducibility of blood pressure variability is moderately high if activity measures are standardised. (Gerin et al., 1993). Schwerdfeger et al, (2014) detail a study utilising CVR and affective states in laboratory and ambulatory contexts, specifically examining if strong reactivity (CVR), and laboratory baselines are related to ambulatory HR. Findings indicate that ambulatory HR and laboratory reactivity were not related, suggesting that laboratory reactivity might not be a reliable indicator of everyday CV activity (Schwerdtfeger, Schienle, Leutgeb, & Rathner, 2014). However, this was mediated by baseline measures, specifically where elevated baseline measures and reactivity were detected elevated ambulatory HR patterns were also present, suggesting that rather than magnitude of reactivity, individuals whose HR follows elevated patterns are more susceptible to ambulatory stressors. In a later study (Schwerdtfeger & Dick, 2019) Schwerdtfeger anchored measurement of HRV to stressful job events in a sample of firefighters, measuring momentary negative affect and resistance. In this instance, RMSSD and HF HRV were associated with stressful events and a lower phasic response to the stressful events was also related to a reduction in reported negative affect.

The difference in the findings in Schwerdtfegers' two studies raise some interesting questions in relation to baseline and negative affect inducing events. Firstly, can the similarities between baselines conditions in HR be expanded to include other more detailed measures of HRV? Secondly, can these be used to examine negative affect inducing events? It should be noted however that the later study still uses a specified negative affect inducing event, emergency calls for fire fighters, that may not be a typically occurring stressor. While these types of study designs do indicate a move towards measurement of ambulatory physiological measures, they are still related to specified tasks that might not be typical for the individual, eg, skydiving or firefighting. Although phasic responses in ambulatory contexts is an important question these tasks may only replicate laboratory type stressors and may not represent a typical or frequently

occurring stressor for the individual (Schwerdtfeger, Schienle, Leutgeb, & Rathner, 2014); which can be further mediated behaviourally, for example by avoidance of certain activities. Another, perhaps more appropriate approach would allow people to self-report negative affect events in their everyday lives that they themselves are more likely to encounter, and this has implications for how measurement timing and framing are specified.

Approaches to analysis of HRV in ambulatory contexts that allow participants to carry out their daily activities with as little intrusion as possible is desirable. If HR measurement is to be integrated using existing technologies, then the goal should be to not alter behaviour to suit measurement but to alter measurement to suit behavioural patterns. Therefore, both HRV measurement and experience sampling should be as un-intrusive as the study design will allow and examine modes of anchoring HR in daily experience. This has implications for physiological measurement, experience sampling and how data that are retrieved from specified timeframes are processed. This could be best achieved by carrying out long-term measurement that allows specificity to shorter term events or contexts. Specifically, twenty-four-hour HR measurements parsed into five-minute epochs that use associated affective and experience sampling to anchor the HR data. This is the focus of Study Four presented here and the proposed methodologies are further outlined in the chapter two. A method of sampling and analysing HRV data is proposed, using predefined contexts on large data sets. These data are used to then examine the associations between a laboratory procedure including baseline and reactivity.

1.6 Experience sampling

In ambulatory studies using physiological measures researchers need to be able to capture accurate and appropriate experiential data. This may include psychosocial variables, affective scales or any number of related state relevant variables. Data capturing experience must be of sufficient granularity to allow the physiological data to be rigorously interrogated but crucially must consider participant burden and other related confounds. Sampling of experience can be conducted using retrospective methodologies or real time sampling and these methodologies are referred to as Experience Sampling Methods (ESM) (Csikszentmihalyi, 2006; Csikszentmihalyi, 2014; deVries, 1992) or Ecological Momentary Assessment (Shiffman et al., 2008; Stone & Shiffman, 1994) (hereafter, EMA). The two terms refer to the same family of measures but have different nomenclature owing to their origins in medicine and psychiatry (EMA) and psychology (ESM) (Walz et al., 2014). Retrospective measures such as the Day Reconstruction Method (DRM) (Kahneman, Krueger, Schkade, et al., 2004) usually sample at the end of day and rely on recall.

Both momentary and retrospective sampling offer advantages and disadvantages and their implementation in study designs depends on the type of data required, with consideration to the population the study is using. Momentary sampling collection methods will allow more accurate association to physiological data however they also carry the potential to introduce experiential confounds. For example, if sampling is carried out too frequently, and the participant experience is altered (for example by annoyance at filling out the measures), (Dockray, O'Neill & Jump., 2019) then is data then truly reflective of experience? Therefore, the type of experience sampling method employed by the researcher needs careful consideration. Another consideration is the research context, for example, whether it requires sampling of data from specific time periods - such as school/home/leisure contexts - or if whole

day aggregate data is sufficient. This is often informed by the type of physiological data sought. For example, in cases where detailed differentiation of specific time periods is sought, timed electronic methods are more desirable. Where whole day aggregate data are required more general contexts, reconstructive methods may be considered. In each case the inclusion/exclusion criteria for each specific measurement period will have to be justified by the researcher and considered for suitability with participants. For example, in a full day reconstructive studies, using aggregate physiological data across the day, data will have to be normalised or excluded, to account for activity or other confounds. Sufficient self-report data will also be required to accurately associate experience/affect with physiological data. Issues exist with all types of sampling that require consideration such as compliance and response patterns in reconstructive studies and compliance in EMA designs.

1.6.1. EMA/ESM

EMA/ESM describe a range of methodologies for conducting research where affective or experimental data are required from participants in ecologically valid, real world contexts. (Shiffman et al., 2008). Sampling strategies aim to capture measures of participants data for specified time-points and provide a within subject representation of participant state for the respective self-report measure. These data are used to assess state in contexts, and/or to derive information about change across time. Data captured by EMA/ESM instrumentation will vary dependant on the research question, for example, researchers can include momentary affect scales or measures related to particular outcomes or health-related behaviours. In the context of the studies presented here, these data provide opportunities for accurate time point associations for combination with biological measures in psychophysiological studies in ambulatory contexts (Myin-Germeys et al., 2018). Developments in technology, and in

particular mobile and wearable devices have enabled a wider range of EMA platforms to be developed and now allow more nuanced and targeted protocols to be implemented.

Proponents of EMA/ESM platforms cite their ability to address issues associated with retrospective designs, such as recall bias, but moreover propose that situationally and temporally sited measurements offer more accurate representations of participant state and also how these may change over time (Burke et al., 2017). While this is undoubtedly the case for certain types of measurement it must be acknowledged that EMA/ESM do carry certain disadvantages, for example, implementation of EMA/ESM platforms can be complex and those that send reminders to participants to complete self-report measures, or instructions to carry out measurements may do so at inconvenient times and can have implications regarding interference with experience and also have implications for adherence (Moeller, et al, 2014). In addition the data yielded from multiple collection periods can be complex and place demand on researchers to analyse (Smyth, 2003). Subsequently retrospective type designs have been included also in research.

1.6.2. Day Reconstruction Method (DRM)

Reconstructive measures are traditionally completed in pen and paper format and rely on recall and as such can carry associated biases and errors. Subsequently, hybrid type measures such as the Day Reconstruction Method (DRM) (Kahneman, Krueger, Schwarz, et al., 2004)have been proposed. The DRM uses a structured daily reconstruction of participant experience in conjunction with specified episodes and affective scales to capture participant experience. The main advantage of the DRM is that it does not disrupt normal daily activities and offers the reflective experience of the diary experience that can produce rich and detailed data that can give participants the opportunity to report salient information that can be missed

by EMA/ESM measures. The DRM also offers a lower possibility of introducing experiential confounds. However time point accuracy and daily adherence to diary completion can be compromised in DRM studies (Diener & Tay, 2014). In addition, the DRM is mostly completed in pen and paper format and is restricted in terms of how participant activity or task completion can be observed by the researcher. This also places restrictions on how the DRM can be used in conjunction with EMA.

This approach of integrative uses of momentary and retrospective type measures is desirable for a number of reasons. For example, on the issue of appraisal, EMA measures capture momentary affect, and while that may be reflective of a particular context and time, if participants are given time to appraise and reflect on an episode it could change over time. For example, a stressful public speaking event at work could be reported as acutely stressful prior to, and during, the event using EMA/ESM but given time afterwards for appraisal could be viewed differently and seen to have some positive benefits. This question speaks to the fundamental principle of whether the experienced self or the remembered self represents a truer reflection of experience (Kahneman & Riis, 2005). While the two concepts have previously been viewed dichotomously, they are now being seen as complimentary and necessary to accurately encapsulate a true picture of experience (Wilhelm & Grossman, 2010a). This represents a more holistic approach to experience sampling and more accurately reflects therapeutic applications for the participant while maintaining the integrity of the data for research purposes. In addition, given that future technological applications using ESM data will undoubtably involve some retrospective interaction with their own data by everyday users, it is important to account for the efficacy of both retrospective and momentary data.

In the studies presented here an adaptation of the Day Reconstruction Method that aims to address a number of these methodological issues while maintaining the integrity of the collection method is presented. Firstly, the DRM is primarily a pen and paper based and does not allow researcher interaction with the data or the participant during the data collection phase. This study hosted the DRM online allowing real time interaction and viewing of data by the researcher during the collection phase. Secondly, with pen and paper DRM studies there is a question of adherence to the specified time for data collection , and patterns of participant adherence based on recorded time of entry of their data are also described. Thirdly, the DRM carries massive resource load in terms of researcher processing of retrieved data. Study three details how the adapted online version reduces this resource load. Finally, it is demonstrated how this collection method has been used in conjunction with physiological data (heart rate) to accurately associate self- report an HRV measures.

Chapter 2: Study one - Measuring the psychobiological correlates of daily experience in adolescents

2.1 Abstract

Mapping the psychobiological correlates of social contexts, experiences and emotional responses of adolescents in their daily lives provides insight into how adolescent wellbeing shapes, and is shaped by, experience. Measures of these psychobiological correlates are enabled by devices and technologies, that must be precise and suitable for adolescent participants. The present review examines the most often used research measures, and suggests strategies for best practice, drawn from practical experience. The rapid advances in technological methods to collect attuned measures of psychological processes, social context and biological function indicate the promise for multi-modal measures in ecological settings. Attaining these methodological goals will support research to secure comprehensive, quality data, and advance the understanding of psychobiological function in ambulatory settings.

Outputs

Published as: Dockray, S., O'Neill, S., & Jump, O. (2019). Measuring the psychobiological correlates of daily experience in adolescents. Journal of Research on Adolescence, 29(3), 595-612. https://doi.org/10.1111/jora.12473

Contributions

Cowritten by all authors: 3.2 Introduction, 3.7 Supporting Participant Adherence, 3.5 Sampling Protocols, 3.18 New horizons, 3.16 Validity of Commercial Applications.

Owen Jump, edited and supported by other authors: 3.3 Ecological Momentary Assessment, 3.9 Measuring Cardiovascular Activity, 3.10 Measuring Heart Rate in Ambulatory Contexts, 3.11 Practical considerations for HR Measurement, 3.14 New Technologies for Psychobiological Measurement, 3.15 Cardiovascular Measurement Devices and Application, 3.17 Recording experience and affect

Siobhan O'Neill, edited and supported by other authors: 3.12 Measuring Sleep, 3.13 Objective and Subjective Measures of Sleep.

Samantha Dockray, edited and supported by other authors: 3. 4 Measuring Salivary Analytes in Ambulatory Settings, 3.6 Saliva Collection Devices, 3.8 Sampling Kits

2.2 Introduction

The psychobiological facets of adolescent experience are essential components in the understanding of adolescent health and development and almost all aspects of the daily lives of adolescents have attracted interest from researchers studying the psychobiology of adolescent wellbeing. Epidemiological studies, as well as smaller scale and prospective, longitudinal studies, have examined the biological correlates of social and behavioural processes, of health behaviours, of stress and resilience, of neighbourhood, school and family context. Mapping how psychobiological processes shape, and are shaped by adolescent experience, emotions and development requires precise and valid measurements of biology in context. There are ever-more rapid advances in technologies and methods that enable psychobiological processes to be measured in ecological settings, and these, combined with developments in conceptual models, have resulted in increasing numbers of research studies including measures of biological processes. Ambulatory assessment methods can measure adolescent behaviour, experience and environment in ecological settings, and there are specific measures and methods that can capture biopsychosocial processes as they occur in these contexts. There are many psychobiological measures that can provide insight into adolescent experiences, although measures using salivary analytes, especially salivary cortisol, and measures of cardiovascular activity, and of sleep are still the most prevalently used in research on adolescence. More recently these psychobiological measures are supplemented with digital and other technologies, including smartphone applications and proprietary devices. However, there are several methodological aspects that researchers using these measures must closely attend to in order to secure high quality data and herein we describe the devices, methods and procedures to measure psychobiological processes in ambulatory settings, along with practical considerations. We also outline established and emerging methods that may be especially suitable for research with adolescent participants in ecological settings.

2.3 Ecological Momentary Assessment

Ambulatory studies that include measures of psychobiology require study designs that can accurately capture the behaviour, affect and experience of the participants. The sampling design and methods should be of sufficient granularity to allow the physiological data to be rigorously examined, and suitable and meaningful measurement periods need thoughtful specification (Burke et al., 2017). Ecological Momentary Assessment (EMA) describes a range of methods used to measure experience (Shiffman, Stone, & Hufford, 2008), and EMA includes real time sampling, such as the Experience Sampling Method (ESM) (Csikszentmihalyi & Larson, 2014) and recall and reconstructive, diary type measures, such as the Day Reconstructive Method (DRM) (Kahneman, Krueger, Schkade, Schwarz, & Stone, 2004) which can provide good estimates of changes of affect and behaviour in the course of the previous day (Dockray, Grant, Stone, Kahneman, Wardle, & Steptoe, 2010). Electronically based Experience Sampling Methods (ESM), including measures that prompt sampling throughout the day, may offer more accurate time point associations to biological measures, but prompts or scheduling of measures may interrupt experiences and behaviour, and contribute to changes in sampling adherence. This may be confounded further with adolescent participants, for example, app based mobile platforms that require immediate compliance to complete self-report measures or which give acoustical reminders or instructions to carry out physiological measurements may do so at inconvenient times or may actually disimprove adherence (e.g. Moeller, et al, 2014). Day Reconstructive Methods (DRM) use structured retrospective daily diaries to reconstruct participant experience and offers a lower possibility of introducing experiential confounds such as interference with naturalistic experience and so is preferable in studies using aggregate data and longer-term ambulatory data. However, time point accuracy and daily adherence to diary completion can be compromised in DRM type studies (Diener & Tay, 2014). With DRM type studies using adolescent participants, we have

found using adapted DRM designs that utilise cloud- based, easily accessible platforms to host survey materials to be useful in addressing time point adherence and accuracy, while also offering novel ways for researchers and participants to interact throughout study periods.

Data captured by all types of EMA instrumentation will vary dependent on the specific research question. In relation to emotion and experience, for example, researchers may want to include momentary affect scales or measures related to particular behavioural outcomes such as eating or sleeping behaviours. When considering the type of assessment to use with adolescents, as with adult cohorts, the type of research question and biological data required should be carefully considered. EMA type designs that use app-based mobile platforms may suit some groups, and offer opportunities when dealing with digital natives, but might be limited in their depth of response. In some cases, the potentially more reflective experience of the paper diary measures may better suit the study design. In the case of adolescent participants, compliance and response patterns may require additional attention. Recent reviews have examined these issues in greater detail (Kim, Kikuchi, & Yamamoto, 2013; Wen, Schneider, Stone, & Spruijt-Metz, 2017), noting that most studies do not report technical or logistical problems, and that adherence to sampling protocol times are still underreported (Heron, Everhart, McHale, & Smyth, 2017). We urge researchers to record and report this data, as precision and completeness of data directly affect the validity of any reported findings.

2.4 Measuring Salivary Analytes in Ambulatory Settings

Saliva has the advantage over other body fluids in that it can be collected safely and easily during regular daily activities, and recent reviews (Gröschl, 2017; Slowey, 2015) have described in detail the collection devices and storage requirements. There are many salivary biomarkers that have been collected in ecological settings and used in studies of adolescent health and development (Granger et al., 2007). Advances in the early 2000s made possible the

processing and assay of these salivary analytes with relative ease and high specificity using enzyme-linked immunoassay procedures (Gröschl, 2017) that are able to provide measures of inflammatory factors such as c-reactive protein, of salivary alpha-amylase, of sex steroids, including testosterone, estrogen and dehydroepiandrosterone. The use of any salivary analyte is driven by the research question at hand however, it would be remiss not to highlight the complex interplay between physiologic systems, and recently Chen, Raine and Granger (2017) provided an example of how the combined, ratio, and interactive effects of several analytes measured in saliva may explain differences in behavioural outcomes more precisely than a single analyte. Although it is possible, and perhaps necessary, to measure more than one salivary biomarker, for the purposes of describing the advantages, cautions and protocols of salivary analytes, here we use salivary cortisol as an exemplar, as it has the most long-standing and evidenced use in adolescent research using measures collected in ecological settings. Salivary cortisol has been used in a number of large-scale, population-based studies of adolescent wellbeing, (e.g. Niarchou, Zammit, & Lewis, 2015; Adam & Kumari, 2009) as well as smaller scale, prospective and longitudinal studies, and its frequency of use in research on human health and behaviour is explained by the significant effect of cortisol on homeostatic mechanisms as its integral role in the stress response (Clow, Hucklebridge, & Smyth, 2018). The collection and interpretation of salivary cortisol data in naturalistic settings presents distinct challenges that represent the challenges of any saliva sampling in ecological setting; here we discuss these and indicate strategies to use in research with adolescents.

Cortisol has a distinct circadian rhythm, rising in the early morning hours, and then gradually declining over the rest of the day to an overnight nadir, completing the 24-hour cycle (Kirschbaum & Hellhammer, 2000). The cortisol awakening response (CAR), occurring 30–45 min after awakening in the morning, is superimposed over the latter phase of this early

morning rise in cortisol and is a distinct component of the diurnal rhythm, and is regulated in a different way (Steptoe & Serwinski, 2016; Stalder et al, 2016). Pertinent to research with adolescents is the difference in the cortisol awakening response that emerges between boys and girls at adolescence, likely explained by reproductive maturation in girls (Oskis, Loveday, Hucklebridge, Thorn, & Clow, 2009) and in models adjusted for sex steroids in adults, women have higher morning cortisol (Juster & Lupien, 2012); separate analyses of correlates of the Cortisol Awakening Response (CAR) for boys and girls is recommended. The phases of diurnal cortisol secretory activity can be quantified by both the dynamics of change and total levels of cortisol secretion with the choice of measure determined by the research question. There is moderate intra-individual stability in the CAR and the decline across the day. The patterns of change in the CAR and in the daily decline have been associated with many psychosocial variables in adolescents, including psychopathology, social relationships, academic performance, health status, including weight status and social contexts (e.g.'s Alink, van IJzendoorn, Bakermans-Kranenburg, Mesman, Juffer, & Koot, 2008; Huynh, Guan, Almeida, McCreath, & Fuligni, 2016; Adam, Quinn, Tavernier, McQuillan, Dahlke, & Gilbert, 2017; Drake, Sladek, & Doane, 2016). Researchers must consider the potential effect of confounding factors such as medication usage, diet and tobacco use (Wilde, Out, Johnson, & Granger, 2013; Nicolson, 2008) as well as non- adherence to sampling protocols (Dockray, et al., 2008; Okun, et al. 2010) and apply exclusion criteria or use statistical processes accordingly. In principle, saliva samples should not be difficult to obtain, but in practice we note several cautions and herein describe strategies to support the collection of high-quality samples.

2.5 Sampling Protocols

The clarity of instruction and emphasis on adherence conveyed to the participants about the need for high adherence to the protocol via participant materials and training sessions is especially important (Stalder et al., 2016). Participants who are convinced of the importance to the integrity of the research in collecting the samples at the designated time are more likely to be adherent (Kudielka, Broderick, & Kirschbaum, 2003). During study induction, the detail and emphasis placed on the sampling protocol is crucial, and it is important to use age-appropriate strategies. These may include using visual materials, including short videos, to show how to use the sampling device, as well as listing or graphically representing the timing of samples. Ethical conduct requires that researchers obtain consent appropriately, and provide information in sufficient detail, but together these may create significant cognitive load, an 'information overload' and this effect may be pronounced for adolescent participants. The training on sample collection must be clear, may use short videos or pictures and include all the instructions, for example before brushing teeth in the morning, or avoiding certain foods and drinks and the many other constraints on sampling timing and integrity (Stalder et al., 2016). In addition to supporting participants to understand and adhere to sampling procedures, researchers must interrogate the literature specific to the analyte they are measuring, as some may be affected by flow rate, or have a response profile to other analytes that may be included in the research, for example salivary alpha-amylase is affected by flow rate (Nagy, et al, 2015; Beltzer, et al, 2010), has a different diurnal (Nater, et al, 2007) and stress response profile to cortisol (Nater et al, 2005) . Researchers can consider separating consent and induction sessions, follow-up session, by phone, or web links to short videos to support participant understanding. Few studies have examined the quality, detail and strategies of providing information on sampling protocol and it is possible that there may be barriers not anticipated by researchers, which affect the return rate for samples collected in ambulatory settings, for

example, information overload, the need to store samples in refrigerators shared with others (e.g. in multiple family households) and other features of the participant group (Adam & Kumari, 2009; O'Campo et al., 2016; Valentino, De Alba, Hibel, Fondren, & McDonnell, 2017).

2.6 Saliva Collection Devices

There are several strategies researchers should consider to increase the likelihood of obtaining quality samples of sufficient quantity for assay procedures. The choice of method (drool or stimulated) and receptacle may be guided by research costs, the purpose of the sample (type of analyte) or volume of saliva required as some strategies may yield volumes insufficient for assay procedures (Granger et al., 2007; Shirtcliff, Granger, Schwartz, & Curran, 2001). Samples can be collected using commercially available devices (e.g., Salivette, Sarstedt) which offer advantages in participant ease of use, as well as in sample processing. The most common strategy involves the use of a small swab to absorb saliva within the persons' mouth, which takes only a few minutes, and then the swab is returned to the collection tube. In commercially available devices, the tube also serves as a storage and transport device and can be centrifuged. This approach is simple and convenient for participants and for researchers. A barrier to using these specific purpose devices may be cost, and so, as devices do not need to be sterile, researchers may elect to use other devices. When research costs are a strong consideration or there are other factors which may prevent the use of commercially available devices, researchers have used other collection devices and strategies, for example Eppendorf tubes or swab-and-syringe procedures (e.g. Carissimi, et al, 2016; Tordjman, et al, 2014) but the decision process must include consideration of the additional processing steps and time; researchers should confirm the required and preferred collection devices with the assaying laboratory before the collection device is confirmed. Another consideration is the ease for

participants. In many collection devices there is a swab that the participant can discreetly place in their mouth, without the swab, participants must allow saliva to collect in their mouth, before transferring it to the tube, usually by allowing a small volume to pool along the inside of their lips and moving it into the tube, using their lips and tongue. This can be awkward and messy, and adolescent participant feedback indicates they will not collect in this way in ecological settings and so simply don't collect samples when they cannot do so in a secluded place, or do not collect the samples at all.

2.7 Supporting Participant Adherence

Until recently, many studies with adolescents used relatively low-technology strategies to remind people to collect a sample, for example, coloured posters or other visual aids for display in the adolescent's home. In several studies where adolescent collection of saliva has been supported by parental research involvement, we have also developed colourful fridge magnets and door signs to remind the parent and the adolescent to collect and store the sample. These strategies were helpful for participants, but do not provide information to determine adherence. With the ubiquity of mobile phones comes the opportunity to use messages, alarms and apps to schedule reminders, as well as to provide checks on sampling adherence. Text-based messaging to mobile phones has been quite successfully used and is a cost-effective method of increasing adherence to ambulatory assessment measurements. Cost-effective automated messaging services may be suitable, although researcher efforts may be equally effective, albeit with higher resource requirements. Noted for high adherence rates, as well as collection of EMA data, a study of salivary cortisol in adolescents in ambulatory settings by Oskis et al. (2009) describes a strategy of sending individualised texts to participants to serve as both a reminder and as data collection, by reply to the text. In our own work we have also used adolescent or parent mobile phones to set alarms, for sampling collection as well as data

responses. Collecting saliva samples in ambulatory settings has both a cognitive load for participants, and may be an undesirable interruption to usual activities, and these demands may be especially pronounced for adolescent participants, contributing to non-adherence to the sampling protocol. There has not been rigorous study of the psychological (e.g., symptoms of depression, self- efficacy), social (influence of peer-participants) and adolescent-specific factors that may affect adherence to sampling protocols, and specifying these factors may contribute to strategies to reduce low adherence. There are also devices and methods to monitor sampling adherence, most often used are contained with lids that record times of opening, for examples, MEMS Trackcaps®. The saliva sampling device is stored inside the bottle or box, and participants are instructed to only open these when collecting a sample; the time stamps provide a measure, albeit indirect, of sampling times. MEMS Trackcaps® are considered to be the highest standard of determining adherence, although high costs are noted (Stalder, et al., 2016). More recently we have trialled the use of smart phones to determine sampling time adherence and collect other data in groups of adolescents and young adults (Dockray et al., 2017). In this method, alarm reminders are set on the participant's smartphone, and these include a web-link to online measures of behaviour, social context and mood. To determine adherence, participants take a time-stamped photograph of the completed sample, showing the barcode or other labelling. The metadata from each photograph can then be extracted and serve as a verification of sampling time. Participants can submit these photos at the completion of the study however, we recommend that the photos are uploaded or emailed immediately, allowing researchers to monitor adherence and provide additional support if needed. Concerns about privacy, data encryption and the use of participant devices may be of concern, and these concerns warrant attention before these methods are more widely used.

2.8 Sampling Kits

There are practical considerations in the preparation of participant materials that researchers should consider when working with participants of any age, but perhaps especially with adolescents and children. The first is the need to have clearly coded sampling times on all collection devices. For researcher use, bar codes or similar (e.g., quick-response (QR) codes) offer efficiency but these are not usually helpful for participants. For participants, we suggest using purpose-specific 'SpitKits', which may include colour-coded tube labels or some other strategy to aid participants in easily recognizing which device they should use for each sample. Many researchers have used colour and numbering systems very helpful to support participant confidence and to reduce sampling errors (Susman et al., 2010; Smyth, Thorn, Hucklebridge, Evans, & Clow, 2015). Attention should also be given to the packaging of sampling devices, especially in multiple sampling day protocols, where we use small plastic resealable bags, each labelled with the participant identification, the day, and other details. In multiple day studies, we have made SpitKits with sampling devices arrayed in tailormade booklets, in order of collection for each day (Matvienko-Sikar & Dockray, 2017). In studies that require participants to collect samples during times when they may be at school or away from home, we provide opaque plastic boxes to conceal the collection tubes, as adolescent participants have described feeling embarrassed, or worried their peers may catch sight of the saliva collection kits and tease them about the samples. Finally, although salivary cortisol has been shown to be stable at room temperature for up to 4 weeks (Kirschbaum & Hellhammer, 2000) and so may be returned by post or be unrefrigerated, for other salivary analytes, researchers will need to consider immediate refrigeration and participant storage arrangements (Nalla, Thomsen, Knudsen, & Frokjaer, 2015; Toone et al., 2013). Not all participants have ready access to a refrigerator, perhaps as they are not at home or related to the social and economic characteristics of the adolescent and their family (O'Campo, et al., 2016) and this may be

addressed by providing cooler bags/boxes. There are several strategies that have been shown to maintain the necessary temperature conditions using cold-packs and cooler-bags provided that time is taken into careful consideration and participants are given clear instructions on handling and storage (e.g. Kruithof, et al, 2014; O'Doherty, et al., 2014).

2.9 Measuring Cardiovascular Activity

Measures of cardiovascular activity, specifically blood pressure, heart rate reactivity, and heart rate variability have been associated with health and wellbeing and so are often used to understand adolescent development and experience. Many studies using ambulatory assessments of cardiovascular activity have used heart rate (beats per minute) and blood pressure, hereafter BP, but the development of portable, non-invasive electrocardiography devices that allow full spectrum measurement of cardiac output now enables researchers to capture a greater range of cardiovascular activity indices used to measure biological response to experience in context. Heart rate variability (HRV) has emerged as a relatively easily obtained measure of responses to- and recovery from daily challenges, as well as positive experiences (Dockray & Steptoe, 2010; Shahrestani, Stewart, Quintana, Hickie, & Guastella, 2015) and can be used as an index of autonomic outflow (ChuDuc, NguyenPhan, & NguyenViet, 2013; Quintana & Heathers, 2014). Typically, time domain measures describe variations in the inter-beat interval (IBI) and spectral analysis of low frequency/high frequency output measures are linked to sympathetic/parasympathetic balance or vagal tone and both measures have psychosocial and affective correlates (Kreibig, 2010; Porges, 2001, 2007; Smith, Thayer, Khalsa, & Lane, 2017; Thayer, Fredrikson, Sollers, & Wager, 2012; Geisler, Kubiak, Siewert, & Weber, 2013). A number of theoretical frameworks, such as the Neurovisceral Integration Model (Smith, Thayer, Khalsa, & Lane, 2017; Thayer, Fredrikson, Sollers, & Wager, 2012) capture the neural correlates of HRV and its' links to affect, emotion,

attention and associated experience and the polyvagal theory (Porges, 2007) outlines the psychosocial correlates such as stress and social engagement (Geisler, Kubiak, Siewert, & Weber, 2013). HR measurement in ambulatory contexts using EMA measures that are informed by these models carry a number of additional considerations with adolescent participants.

2.10 Measuring Heart Rate in Ambulatory Contexts

In general, adolescents experiencing social stressors have larger reductions in HRV than adult participants, a trend that may be related to reduced autonomic reactivity and better coping strategies associated with aging (Shahrestani et al., 2015). Researchers using HRV with adolescent participants should consider this in the interpretation of data and be cognizant that there is debate regarding the use of these measures as dependent variables or in correlative studies (Billman, 2013; Reyes del Paso, Langewitz, Mulder, van Roon, & Duschek, 2013), and that age, gender and body weight (BMI) may further influence heart rate variability and pubertal stage has been shown to influence HRV as well as other physiological parameters in adolescents (Komkova, Ermakova, & Selverova, 2017). Researchers intending to use laboratory stressors or cognitive tasks with adolescents should be aware that stress reactivity may correlate with pubertal stage or pubertal tempo (Romeo, 2017; Stroud, Papandonatos, Williamson, & Dahl, 2011; Saxbe, Negriff, Susman, & Trickett, 2015). There are other, additional considerations when designing methodologies for use with adolescent participants. The neurovisceral integration model posits prefrontal, top down vagally mediated control of emotional states that are captured in HRV measures, and the cortical development that occurs during adolescence, correlated with chronological age and developmental stage may influence their ability to regulate affect and so affect HRV parameters (Koenig et al., 2018; Antelmi et al., 2004; Koenig & Thayer, 2016; Silvetti, Drago, & Ragonese, 2001). Studies in this area

must also consider the familial and social context of adolescent participants. Parental factors, such as socialisation (Williams & Woodruff-Borden, 2015) or psychopathology may influence emotion regulation and resulting physiological parameters in adolescents. Therefore, the social milieu adolescents inhabit may need to be adequately measured using additional parental measures of psychological and social functioning to accounts for confounds in the analyses of cardiovascular measures. If designated time periods or events are of interest, for example, exam periods, parental interaction or social engagements, researchers may use the margins at these periods as anchor-points to examine heart rate variability between designated periods. These will be determined by the type of EMA used by in the study design. In studies using HR measures researchers will have to detail how they account for confounds such as movement and these confounds will have to be interpreted in light of the EMA data and assessed for inclusion or exclusion in final analysis. It is worth noting that both reconstructive and EMA methods are sufficient for use in HR studies if accelerometer and timing data is available from the measurement device. Another consideration is whether the data must be sampled from specific time periods- such as school/home/leisure contexts- or if whole day aggregate data is required. Finally, comparisons of EMA and DRM based measures (Diener & Tay, 2014; Dockray, Stone, & Steptoe, 2011) show sufficient agreement to warrant inclusion of either technique. If detailed differentiation of specific time periods is required, then timed electronic methods are necessary, where aggregate data are used, or comparison of context is the focus, the DRM is recommended. In each case the inclusion/exclusion criteria for each specific biological measurement period will have to be justified by the researcher and considered for suitability with adolescent participants. For example, in a full day reconstructive study, using aggregate physiological data across the day, data will have to be normalised or excluded, to account for movement or other confounds (Laborde, Mosley, & Thayer, 2017).

2.11 Practical considerations for HR Measurement

In our experience there are several practical considerations of using heart rate variability or other measures of cardiovascular function in ambulatory contexts with adolescents. Many devices are waterproof or resistant and can collect data for up to three weeks, depending on the measurement settings programmed by the researcher, and there are opportunities to collect data over 2-3 weeks with few problems. Researchers must ensure the devices are suitable for use with adolescents, as the algorithms used to process the data, or the sizing or weight of devices developed for use in adults may not be suitable for adolescents, who may not only have a smaller body size than the equipment is designed for but may have differences in parameters of physiological function. Devices must be calibrated, researchers must follow a standardized approach to fitting and maintaining the equipment, as well as procedural methods to minimise measurement errors. Some devices, for example Actihearts (Kristiansen et al., 2011; Takken et al., 2010) used to measure heart rate variability, can be very small but they are still obvious on the body. Some devices make audible noise, or have a flashing light to indicate function, or are attached with a larger adhesive patch or strap, and so where participants are concerned about the appearance or noise of the device, researchers may need to accommodate participant requests to remove the device for certain events, or risk unexplained data loss, or refusal. In these cases, researchers may need to reschedule the data collection period, or arrange a follow-up visit to re-fit the device. Researchers must be well trained to fit the device to ensure a clean and strong signal from the cardiovascular or other physiological system is detected or fitted so that the device is comfortable to wear, and this training must be specific to adolescent participants. There are also challenges of fitting the device on female torsos (Rautaharju, Park, Rautaharju, & Crow, 1998) with participants with developing or larger breasts, or where there is chest hair, which may present both challenges in the technique for device fitting, and to participant comfort. Local ethical requirements may require two researchers to be present when

doing research with adolescent participants and this itself may add to participant discomfort as the participant may need to partially undress. These issues specific to the placement and wearing of devices on the adolescent participant may affect recruitment, compliance and attrition, and we urge researchers to develop strategies to increase the acceptability of cardiovascular activity measurement devices to adolescent participants. Recent advances in the technology of remote and wearable technologies suggest the imminent possibility of continuous, remote and real-time monitoring of electrocardiogram, heart rate, heart rate variability, respiration rate and skin temperature, embedded in clothing (e.g. Salvo, Pingitore, Barnini and Di Framncesco, 2017; Bandodkar & Wang, 2014), however extensive validation work remains to be completed before these may be acceptable for research purposes.

2.12 Measuring Sleep

There are changes in the homeostatic and circadian regulation of sleep that are typical of adolescence, usually associated with shorter sleep duration, later bedtime and later or earlier wake times, and later wake times on non-school days (Olds, Blunden, Petkov, & Forchino, 2010; Hagenauer, Perryman, Lee, & Carskadon, 2009). The sleep-wake cycle has an essential role in basic functioning and during adolescence there are changes in sleep patterns that often result in more “owl -like” patterns of sleep (Carskadon, Vieira, & Acebo, 1993; Zeiders, Doane, & Adam, 2011) and Roennenberg et al. (2004) have previously proposed that changes in sleep/wake preference may serve as a marker of the end of adolescent period. There are four areas in the biology of sleep regulation that show changes during adolescent development. These are: (1) a decrease in the duration of non-REM and REM sleep, (2) the development of a more adult like pattern of REM sleep, (3) an increase in daytime sleepiness, and (4) there is a shift in circadian rhythms to a more evening type preference (Dahl & Lewin, 2002). This change is accepted to be a result of biology, changes in melatonin secretion (Crowley et al.,

2007), interacting with multiple social and other environmental factors such as social interactions, night-time activities, and stress and psychosocial wellbeing (Gamble et al., 2014; Gregory & Sadeh, 2012; Moore & Meltzer, 2008). The processes of puberty contribute to the biological shift towards later bedtimes in adolescents, and this phase-delay is correlated with maturational stage. Accurately mapping patterns of sleep behaviour are essential to describing how sleep is related to health and wellbeing, especially as sleep timing, quality, preference and duration may anchor other psychobiological processes, or may mediate psychobiological pathways, such as diurnal rhythms of cortisol (e.g Ly, McGrath, & Gouin, 2015; Susman, et al 2007; Fuligni and Hardway, 2006; Fuligni, Arruda, Krull, & Gonzales, 2018) and responses to daily experiences, including stress (Mrug, Tyson, Turan, & Granger, 2016) and health behaviours (e.g. Dolsen, Wyatt, & Harvey, 2018). Overall quality of sleep during adolescence is lower than in childhood (Crowley, Acebo & Carskadon, 2007) and this can have profound effects on wellbeing, not only on the following day, but on general psychosocial wellbeing, including risk-taking behaviour (O'Brien & Mindell, 2005; Pasch, Laska, Lytle, & Moe, 2010), depression (Lovato & Gradisar, 2014), anxiety (McMakin & Alfano, 2015) and academic outcomes (Fuligni, Arruda, Krull, & Gonzales, 2018). Many of the studies of adolescent sleep and biopsychosocial correlates and consequences use self-report data collected with sleep behaviour diaries, the gold standard for subjective sleep assessment, but many nights of data are required for accuracy. In an illustration of the need for multiple days of diary data, Short, Arora, Gradisar, Taheri, and Carskadon (2017) report that at least 5 nights of data may be necessary to accurately measure sleep onset latency, bedtime and sleep duration, and the number of measures may vary across school and non-school day diaries, and social and cultural context. A recommendation for this study also is that daily download and use of electronic diaries would avert the prospective or delayed completion of diaries by adolescent participants. To describe the interplay of biology, behaviour and context in sleep onset and duration, we

suggest the use of both objective and self-reported measures of sleep, with assessments conducted in real-time, not least to minimize recall bias.

2.13 Objective and Subjective Measures of Sleep

There are many ways to measure subjective sleep. Methods include self-report diaries, charts and questionnaires, which can be completed online or in paper, or other methods including phone calls, have been used by researchers interested in adolescent sleep, including standard and study specific measures of sleep quality and quantity. Measures of sleep quality and quantity can be collected using diaries and questionnaires, and also objectively using measures of biology including Electroencephalography, polysomnography and actigraphy. Historically, questionnaires and diaries have been used to collect sleep data (Libman, Fichten, Bailes, & Amsel, 2000; Rogers, Caruso, & Aldrich, 1993), however, in recent years objective measures of sleep are frequently used, such as actigraphy, and self-report EMAs. Actigraphy is a non-invasive method of measuring activity and rest levels. These devices are worn continuously (usually on the wrist, (e.g. Actigraph) and detect body movements, and others (e.g. Actiheart) detect heart rate variability and movement, these recordings can be used to detect sleep parameters including performing sleep stage analyses (Jackowska, Dockray, Hendrickx, & Steptoe, 2011). Concurrent assessment of reported and actual sleep behaviour, and behaviour that affects sleep, provides insight into the changing sleep patterns associated with the adolescent period, however these methods may not be concordant. For example, actigraphy may overestimate the total duration of sleep and mis-classify periods of motor activity during sleep as periods of wakefulness (Lockley, Skene, & Arendt, 1999) and adolescents may under- or over-report their sleep, and this may be related to total sleep duration. For example, Lauderdale, Knutson, Yan, Liu, and Rathouz (2008) indicated that participants who slept for 5 and 7 hours over-reported sleep duration by 1.3 and 0.3 hours respectively. The moderate

correlation of objectively and subjectively measured sleep, and the reliability and validity of self-report of sleep represents a current challenge, as most studies of adolescent sleep are done in ambulatory settings, rather than in sleep laboratories (Sadeh, 2011), and further work to determine the concordance of actual, reported and measured sleep, as well as longitudinal studies of adolescent sleep changes using objective measures of sleep would provide a surety about measurement methods, as well as map the patterns of sleep as related to changes in age, in biology and in psychosocial development.

Improved accuracy of both objective and subjective measures of sleep is possible with the use of an integrated data collection method between devices and EMA method, specifically the use of smartphone applications (e.g. Bianchi, 2015; Sano et al., 2015). Smartphone apps may be preferred over paper diaries by adolescent participants, and so using smartphone technology presents a new opportunity to enhance data collection and precision, at relatively low cost. However, reliable, precise actigraphic smartphone applications are yet to be validated in adolescents (Bhat et al., 2015; Behar, Roebuck, Domingos, Gederi & Clifford, 2013). A challenge to using EMA data collection in adolescence, in any format, is the event contingent nature of the data collection. Societal structure for adolescents is much more rigid than for adult participants (Shaw, Caldwell, & Kleiber, 1996) and therefore EMA requirements for recording data at event-contingent or random time-points may not be feasible, for example during school hours. However, with the use of smartphone apps which send prompt reminder to participants there is potential to offset this issue by allowing participants to specify periods of suspension from prompting during times of unavailability. Researchers may opt to use text messages to encourage adherence for objective measures (e.g. actigraphy) as well as the collection of EMA data (Oskis et al., 2009; Schnall et al., 2013) and these can have low participant burden (although high intrusion), higher efficacy when combined with other

objective assessments, and the ability to send automated reminders (either by in-app notifications or text messaging) (Dunton, Liao, Intille, Spruijt-Metz, & Pentz, 2011; Garcia et al.., 2014; Runyan et al., 2013).

2.14 New Technologies for Psychobiological Measurements

Advances in technologies have enabled researchers to collect vast and often rich data on adolescent mood and behaviour in ecological settings (Brannon, Cushing, Crick, & Mitchell, 2016; Duval, Fujisawa & Hashizume, 2007; Yang & Hsu, 2010). Wearable sensors make it possible to assess the dynamic activity of physiological rhythms and function in ways and settings that were previously impossible, or at least impractical, and there are now research-quality sensors to measure physiological processes, including skin conductivity, heart rate, respiration, blood pressure, electrocardiography and electroencephalography. These, in combination with technologies and apps that collect, collate, and transmit the data have increased precision and reduced the need for participants to upload data, are creating new measurement possibilities (Trull & Ebner-Priemer, 2009). Smartphones are already used as devices for data collection, including for measures of biology such as pulse volume and pulse rate variability (e.g., Heathers, 2013; Matsumura & Yamakoshi, 2013), sleep (Chen et al., 2013) and other psychobiological measures (e.g., Tran & White, 2012; Rao, Hou, Golnik, Flaherty & Vu, 2010).

2.15 Cardiovascular Measurement Devices and Application

Researchers should be cautious in adopting app-based measures of physiological processes without verification that the data are valid and reliable. Consumer market apps have been shown to provide highly inaccurate data (e.g. Alexander, Minhajuddin & Joshi, 2017; Plante et al., 2016) and are currently not recommended for research purposes. However, driven by

advances in patient monitoring and medical technologies, it is likely that app-based measures of psychobiology in ambulatory settings will advance at pace and the recent advances in ambulatory assessments of steroid hormones using lateral flow devices are an example (Miocevic et al., 2017; Shirtcliff et al., 2015). In some studies, it may be necessary to develop bespoke applications for data collection, however, translating researcher needs and ideas into app development is not always feasible from a technological or data security perspective (Heron, Everhart, McHale, & Smyth, 2017), and engaging third parties to participate in the development of the application has significant cost implications. Symbiotic partnerships with computer science research groups, for example, may enable research on the psychobiological experience of adolescents, although currently the emphasis is on psychosocial health and experience (e.g. Magallón-Neri, Kirchner-Nebot, Forns-Santacana, Calderón, & Planellas, 2016; Drake, Sladek, & Doane, 2016).

2.16 Validity of Commercial Applications

Many devices used to measure ambulatory physiology offer practical solutions but, researchers may not have full knowledge of how the data are processed as a result of proprietary software restrictions. For example, opensource software that allows detailed analysis of each successive heartbeat is available but requires researcher expertise for detailed analysis (Ellis, Zhu, Koenig, Thayer, & Wang, 2015), in contrast consumer devices such as ‘smartwatches’ cede little or no information about how data are processed and what algorithms are used, including how irregularities such as missing data and anomalous readings are processed. Researchers must ensure that collection devices, as well as any algorithms or post-collection data processing, are reliable and valid, but we highlight the significant potential for the use of commercially available equipment in research studies. There are also opportunities in the capacity for smartphones to collect video and sound recordings, take pictures, provide Global

Positioning System (GPS) data and social environment information, including social density and social engagement, and to monitor mobile-app usage (Heron, Everhart, McHale, & Smyth, 2017) – all potentially meaningful in relation to describing how psychobiological processes may be reactive to, and recover from, experiences. The validity of the psychobiological measures that can be obtained using smartphones, sensors, (including all types of wearables) are only one aspect researchers may need to be concerned with, as there may be unique, and heretofore less considered, ethical and legal considerations of using smartphone and other devices for research (King, 2011). Adolescents may be especially concerned with entering information in ambulatory settings, as they are more likely to be 'shoulder surfed' (Muslukhov, Boshmaf, Kuo, Lester, & Beznosov, 2013). There may also be concerns specific to some behaviours, for example illicit behaviours or access to data by parents or other adults, (Capon, Hall, Fry & Carter, 2016; Hang, Von Zezschwitz, De Luca, & Hussmann, 2012) as adolescents from some contexts are more likely to share access to mobile technologies with family members (Yardi & Bruckman, 2012).

Smartphone apps have made it increasingly easier to collect EMA data from participants in high detail and participants can respond with relative ease and low burden. Participants can be alerted via apps as to when they need to make an entry for time sensitive information and other sampling collections, for example when collecting saliva samples. Many research contexts or populations have mobile phone saturation, so the likelihood of a barrier to participation is much reduced than in the past (Lauricella, Cingel, Blackwell, Wartella, & Conway, 2014) although in some circumstances, researchers may need to supply an internet-enabled device and internet access.

2.17 Recording experience and affect

The aim of most research using psychobiological measures of the person in ambulatory settings is to relate the experience to biological processes, and so biological measures must be related to measures of affect and experience. Reports of daily experiences have often been collected with paper- and phone-based diaries, although these have high participant and/or researcher burden (Shiffman, Stone, & Hufford, 2008; Stone, Kessler, & Haythomthwatte, 1991). Text messaging diaries can be effective for collecting brief, structured diary data, especially on sensitive topics, for example, substance use and smoking (Phillips, Phillips, Lalonde, & Dykema, 2014; Schober et al., 2015), and there are specific programmes to do so including the Youth Ecological Momentary Assessment System (Garcia, et al., 2014). Social media programmes with private account functionality may be acceptable to adolescents, for example some researchers have utilised Twitter to collect EMA data on experiences of stress but we flag concerns about data protection and ownership. Both intermittent and asynchronous logging of experience is possible and useful, for example, with saliva sampling, when timing precision may be essential an asynchronous collection of EMA data is best, however, for other measures such as cardiovascular activity data, intermittent logging may be more suitable. There may be difference in burden and so adherence between asynchronous and intermittent data collection for psychobiological analysis, although both may be used in a single study.

We recently studied stress and sleep in adolescents using paper-based diaries with both asynchronous and intermittent logging to collect data on daily experiences of stress, daily emotions, sleep patterns, food intake, and saliva sampling (O'Neill & Dockray, 2014, 2015). Participants completed diaries for 4 days each, every 6 months, for 12 months; the first diary (Sampling Diary) was an asynchronous record of saliva collection, with text messaging at each time point as a reminder to collect a sample and complete the diary. A second diary of daily

hassles and uplifts required both intermittent and asynchronous logging and participants were asked to complete the diary at least once at the end of each day, but preferably at three time points; morning, afternoon, and evening. Notably, completion rates were much higher for the Sampling Diary than the hassles and uplifts diary. We attribute this to the text-message prompts, and simple response format, especially as in the 'daily hassles and uplifts' diary questions that required more effort had lower completion rates than lower burden questions. Whatever type of logging is used, strategies to encourage completion may be useful, including financial compensation, unique rewards (e.g. genotypic or sleep profiles), or novel lottery techniques (Burke et al.,2017; Goldschmidt et al., 2014; Martinson et al., 2000) including where the chances of a reward are increased with increasing numbers of diary completions (Burke et al., 2017).

2.18 New horizons

Ambulatory assessments of psychobiological function, using biomarkers in saliva and cardiovascular system activity, enable the mapping of how the activities, emotions and contexts of daily life alter biological function. The development of devices, methods, and technologies to accurately capture biological measures has advanced rapidly over the last decade, with a concomitant rise in the number of studies of the psychobiology of adolescent experience. Many devices have been adapted or developed for ambulatory assessment of adolescent experience, and these can return rich datasets. However, the integrity and validity of these datasets depends on well-considered theoretical models as much as it depends on the conduct of the practical aspects of the research, which must be developed with specific thought to working with adolescent participants. We have noted the need to develop approaches to collecting measures of psychobiological function in ambulatory settings that are specifically developed to suit adolescents, and this is crucial if ecological validity is to be attained. The potential of

ambulatory assessments is directly linked to the innovations in measurement devices, including in smart-fabrics and instruments, and smartphone

apps to collect psychosocial measures, and we anticipate the next wave of innovation will support concurrent recordings of context, mood, behaviour and psychobiological function and highlight the need for advancements in statistical analytic approaches to fully describe the association of experience, emotion and psychobiological processes in ambulatory settings.

Chapter 3: Methodology

3.1 Methodology overview

The collected methodologies used in the studies presented in this PhD were part of a combined data collection phase, designed to run concurrently, including the laboratory study and the four-day ambulatory study. Although the methodology for each study is presented in its respective chapter, this section provides additional information relevant to the context of the PhD, including: the selection and recruitment of participants, the full range of materials and measures employed, the study procedures, data analysis, the pilot studies that were used to refine procedures, and ethics. The studies were carried out as follows,

Study One (Chapter two): Study one is a summary narrative review titled, *Measuring the psychobiological correlates of daily experience in adolescents* (Dockray, O'Neill & Jump., 2019). This work details the BEaTS laboratory group's work focussing on mapping the psychobiological correlates of social contexts, experiences and emotional responses, with a particular focus on adolescent participants. The review provides insight into how adolescent wellbeing relates to psychobiological correlates and how psychophysical and experience sampling methods are enabled by available devices and technologies. The review examines the current literature in the field, frequently used research methods, and suggests strategies for best practice in conducting research in this area drawn from the practical experience of the research group. This also provides the context within which the PhD was conducted, with emphasis on the comparison of laboratory and ambulatory measures. The data collected, and the methodologies applied, in the overall PhD studies were used to inform the review presented in chapter three.

Study Two (Chapter four): Study two, *Cardiovascular responses to stress utilising anticipatory singing task* (Jump & Dockray., 2020), is a laboratory-based study that uses a within-subjects design to evaluate the viability of an anticipatory singing task to elicit a physiological response utilising cardiovascular measures. The task includes two stressor components, a standardised math task followed by an anticipatory 'sing-a-song' task in Resting/Reactivity/Recovery segments that adhere to measurement standards and practice for HRV measurement. Patterns of tonic and phasic HRV are examined to assess the efficacy of the two task components, separately, in combination, and, in comparison to baseline and recovery phases. This study is included because in order to compare laboratory measures to ambulatory contexts a standardised laboratory measure was required, and this presented an opportunity to examine a novel task stressor component, the anticipation of a singing task. This study includes the novel task, along with the standardised maths task, and sought to identify if the novel stressor prompt would differ from baseline measures and compare the magnitude of response to the established maths task.

Study Three (Chapter five): *Examining the use of an online adaptation of the Day Reconstruction Method* (Jump & Dockray., 2020), uses a between-within subjects' design to evaluate an adaptation of the Day Reconstruction Method (DRM). In this study original data were collected over a four-day measurement period using an electronic version of the DRM that was adapted specifically to enable online data collection during the ambulatory study. During the planning and design phase of the ambulatory study the it was noted that three major issues existed with the DRM tools (the literature and rationale for this are detailed in the DRM paper) that existed previously. Firstly, there was no way to track adherence, secondly the pen and paper format was laborious for both researchers and participants, and finally, with pen and paper formats integration of the DRM with other experience sampling methods was difficult.

Study three presented here examines those challenges using novel data and compares our findings to existing DRM studies.

Study Four (Chapter six): “From the lab to the living room: Measuring heart rate variability in ecologically valid contexts”, uses a within-subjects design to compare a laboratory-based stress task, including a standardized maths task followed by an anticipatory ‘sing-a-song’ task in resting / reactivity segments, and patterns of HRV in matched ambulatory contexts across 4 x 24-hour periods for each participant. This study is the final study in the PhD and represents the culmination of the work that was conducted in the laboratory and in ecologically valid contexts. The methodology in this study represents a novel means of data collection using HRV measures in ambulatory settings. The sampling method captures four full days of uninterrupted ECG measurement. The ECG data are then sampled using the DRM data described above to extract epochs that are defined by participant experience. These data are then compared to the laboratory measures and the associations and differences are described. This design was decided upon in order to contribute to work that links laboratory and ambulatory measures. Study four further details the relevant literature that describes the rationale for using this methodology.

3.2 Participants

3.2.1 Recruitment

Participation for Studies Two, Three and Four, was sought via advertisements on social media (Facebook and Twitter), and university email. Once an initial expression of interest was registered, the participants were contacted via email. Participants were then sent a link for initial surveys to complete and to self-assign to time for laboratory attendance. Initial surveys resulted in 84 expressions of interest. Participants were asked to self-assign to the study they would like to participate in, either the full four-day ambulatory study or the laboratory study alone or both.

The data was examined to assess participants' self-reported medical conditions or medications, including, cardioactive medications, antidepressants or anti-hypertensives (Cohen, 2001; Kemp et al., 2010; Schroeder et al., 2003) that would preclude them from participating. Participants also reported data on physiological and lifestyle factors (both stable and transient), prior to attending the laboratory related to the measurement of heart rate data in accordance with recommendations set out in Laborde et al. (2017), including, drug and alcohol use (Daniel S. Quintana et al., 2013), and a brief screening tool for depression and anxiety, the PHQ-4 (Kroenke et al., 2009). In addition participants also self-reported age (Umetani et al., 1998), gender (Koenig & Thayer, 2016). Each of these factors is detailed further in the materials and measures section of studies two and four.

Following the initial survey phase, the participants were assigned to each respective study resulting in the following numbers, for study two 45 participants, for study three 29 participants were included, and for study four 25 participants completed the full four-day study. The inclusion/exclusion criteria and final numbers are further detailed in each study respectively.

3.3 Materials and measures

The following section gives an overview of the materials and measures included in the laboratory and ambulatory studies. Further detail, including the justifications for their inclusion are detailed in each study respectively.

3.3.1 Initial survey phase.

Recruitment for the laboratory and ambulatory studies was carried out in two phases. Phase one sought to recruit and determine participant suitability from the larger general population. In order to identify participant eligibility an initial survey was conducted that included a number of questions and scales, including; demographics, drug and alcohol use - *Simple Screening Instrument for Substance Abuse* (Center Substance Abuse Treatment, 1994), depression and anxiety - *An Ultra-Brief Screening Scale for Anxiety and Depression PHQ-4* (Kroenke et al., 2009). In addition, participants were asked to self-report any medications such as psychoactive or cardiovascular medications. The decision was taken to seek this information at initial phase to identify any physiological factors, such as medications (Kemp et al., 2010) that would preclude participants. Secondly, psychological factors such as anxiety and depression have been studied in HRV type studies, (L. Brown et al., 2018; Chalmers et al., 2014) and these responses were included later in data analysis, however no data were excluded.

3.3.2 Demographics

Demographic data were sought regarding age, gender, employment, income, marital status. Demographic data was collected at the initial survey phase, prior to the laboratory and ambulatory phases. The demographic measures were informed by the data required for the Day Reconstruction Method (DRM) (Kahneman, Krueger, Schwarz, et al., 2004).

3.3.3 Physical data

Physical characteristics of the participants were recorded in the laboratory, including, mass (kg), height (cm), (Antelmi et al., 2004). Participants were also asked to self-report age and gender and estimated resting and maximum active heart rate. All of the above parameters are required by the Actiheart software as to account for confounds and group differences in HR profiles, such as those correlated with Body Mass Index (BMI). BMI is automatically calculated using the Actiheart proprietary software.

3.3.4 Electrocardiogram measures

ECG measurement for studies two and four was carried out using the Actiheart measurement device (Brage, Brage, Franks, Ekelund, & Wareham, 2005). The Actiheart has been validated by comparison with standard clinical measurement devices (Holter ™)(Kristiansen et al., 2011) and is a non-invasive method of measuring physical activity and energy expenditure (PAEE). (Crouter et al., 2008). The Actiheart device is attached over the skin and can be used for long recording periods (up to 21 days) in an ambulatory setting. ECG is measured with a sensitivity of 250uV and a range of 30 bpm – 250bpm. ECG is detected, amplified and sampled at 128Hz and the device is capable of capturing and calculating full r-r wave intervals necessary for HRV derivations. Actiheart uses the resultant data in its proprietary software to calculate and resolve HRV measures that are presented to the user as follows; the r wave negative edge is detected using the Pan Tompkins QRS detection algorithm (Pan & Tompkins, 1985). Actiheart returns ECG data for defined epochs and its proprietary algorithm sorts and cleans data prior to presenting it to the user. R-R intervals for at least 16 heart beats are captured for each epoch and averaged (CamNtech, 2018). Values outside of a +/- 25% range are removed and the remaining data re-averaged. This represents the data returned to the user. Additional quality indices (range 0-1) for each epoch are returned and can be examined by the user and excluded

if necessary. For the purposes of the studies that use ECG measurement here an additional quality control procedure at the data analysis phase (Table 3.2) was included. Any epoch with a sub optimal quality rating was excluded from the analysis. This was determined by examining the quality percentage returned by Actiheart and the exclusion of any data with a quality rating of below 90%. Further detail of how these criteria were applied is outlined in the Signal Testing section 2.3.8 and the Data Analysis section of this methodology.

3.3.5 Physical activity

The Actiheart device returns measures of physical activity for the participant as movement in 'counts of unit per time' at an 8-bit resolution in a frequency range of 1hz-7hza and sampling rate of 32Hz. The device contains a piezo electric element that generates a measurable transient charge when subjected to acceleration. The resultant charge produces a voltage signal that is divided into a predefined range of positive and negative levels (- 128 to + 128). This range is sampled 32 times a second. This time/second measurement is then applied by the proprietary software to the user defined epoch (CamNtech, 2010, 2018) returning the activity measure for the epoch. It should be noted that the manufactures do not specify how movement data relates to SI units such as m/s. The inclusion of a device capable of measuring movement was required for both the laboratory and ambulatory phases. This was one of the key inclusion/exclusion criteria for Study Two and in particular the ambulatory phase of Study Four for epochs reported. For all studies presented here, including the ambulatory phase, only epochs with NO movement were included in the analysis.

The combined measures captured by Actiheart represent an overview of the participants' movement and heart rate during a specified epoch. Information is downloaded via accompanying proprietary computer software to allow examination of heart rate. A summary of the full range of measures is presented in table 3.1

Table 3.1: *Range of Actiheart HRV Measures*

	Measure	**Description**
Reactivity	BPM (Beats Per Minute)	Beats Per Minute included as an index of overall reactivity
Time domain	RMSSD (Root mean square of successive differences)	Root Mean Sum of Square differences
	IBI (Interbeat/R-R Interval) IBI max IBI min Mean IBI	Interbeat interval- measure of Max/min/and average IBI were taken
Frequency domain	LF (Low frequency)	Low frequency component of heart rate
	HF (High frequency)	High frequency component of heart rate
	LF/HF (Low/High frequency ratio)	Ratio of Low to High
Signal quality	Range 0-1 Including excluded measures	Actiheart proprietary measure of signal quality
Physical activity	Range +128 to -128 Minimal/Moderate/Vigorous (Researcher defined activity)	Actiheart proprietary measure of physical activity

Self-report measures

3.3.6 The Ultra-Brief Screening Scale for Anxiety and Depression

PHQ-4 (Kroenke et al., 2009).The PHQ-4 is a brief screening tool used to assess the presence of symptoms of depression and anxiety. The scale is presented to participants as a 4 item self-report scale. Responses are given on each item on a 4-point Likert scale ranging from 0-3. Participant responses return scores on two factors, depression and anxiety. Scores ranging from 0-12 indicate severity of depressive and anxiety symptoms. Scores are broken down as:0-2, normal; 3-5, mild; 6-8, moderate; and 9+, severe. The scale has been previously shown by Kroenke to have a good internal reliability of α-.80 (Kroenke et al., 2009). The PHQ -4, as with similar brief scales such as the PHQ-9, was developed as a method of quickly assessing patient depression and anxiety in a clinical setting. The scale has been used across multiple research contexts, both as indicative scales and screening tools for comorbidity/confounds

within samples. The PHQ-4 has been included in this study to allow for brief screening for depression (Brown et al., 2018; Kemp et al., 2010) and anxiety (Chalmers et al., 2014; Levine et al., 2016) as confounding factors.

3.3.7 Simple Screening Instrument for Substance Abuse

The *SSIS screening tool* (Center Substance Abuse Treatment, 1994) is an instrument for detecting substance abuse. It is a 16-item scale (14 of which are scored, Items 1& 15 are lifestyle indicative questions but are not included in final scoring) designed to detect symptoms of alcohol and substance abuse in the past six months. Participants return scores on each of the 16 items across five domains in a mixed response format. Questions are presented in a mixed format as dichotomous yes/no while others have a multiple choice or check box format. The scale has been shown to have a good internal reliability of .85 (Peters et al., 2000) The SSI-SA consist of five domains: substance consumption, which measures frequency, length and amount of use; preoccupation and loss of control, which examines amount of time spent concerned with word(s) missing; adverse consequences; problem recognition; and tolerance and withdrawal (Center Substance Abuse Treatment, 1994) The scale was included as a screening tool to account for possible confounding effects (Daniel S. Quintana et al., 2013) caused by drug or alcohol use on heart rate data prior to participation in the study. Anything on validity?

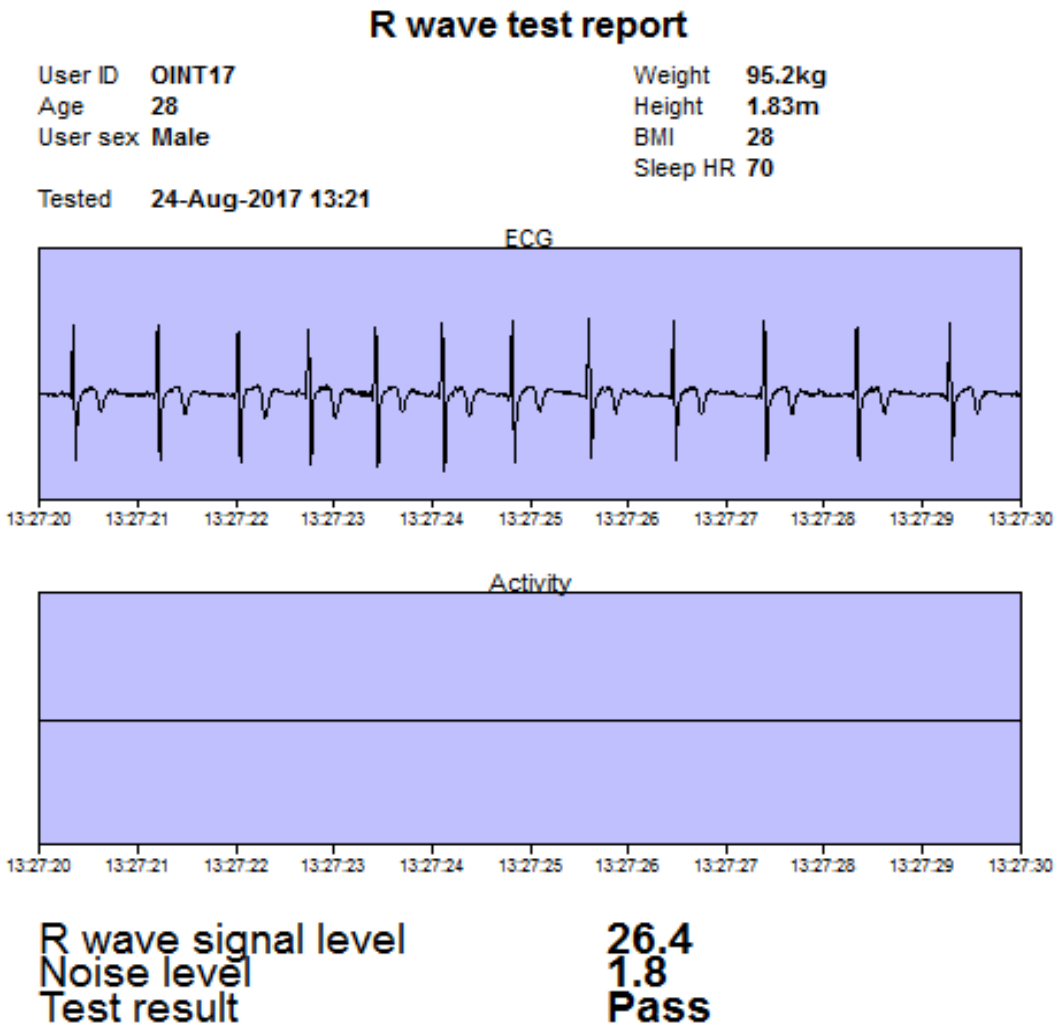
3.3.8 Actiheart signal test measures

Each participant was required to complete a standardised signal test to ensure the quality of the ECG recording. The signal test is pre-defined by the manufacturers of the Actiheart device as a ten-minute measurement period that returns heart rate and movement data. The signal test facility allows detailed examination of each recorded test epoch. The data is presented, as shown in Figure 3.1 to the researcher and an overall pass/fail for the R-R signal quality and noise level is returned. A wave test report is also produced Figure 3.2 and noisy epochs or ones

where the signal is lost are presented to the user to allow decisions regarding the inclusion/exclusion process.

Figure 3.1: *Actiheart signal test user interface*

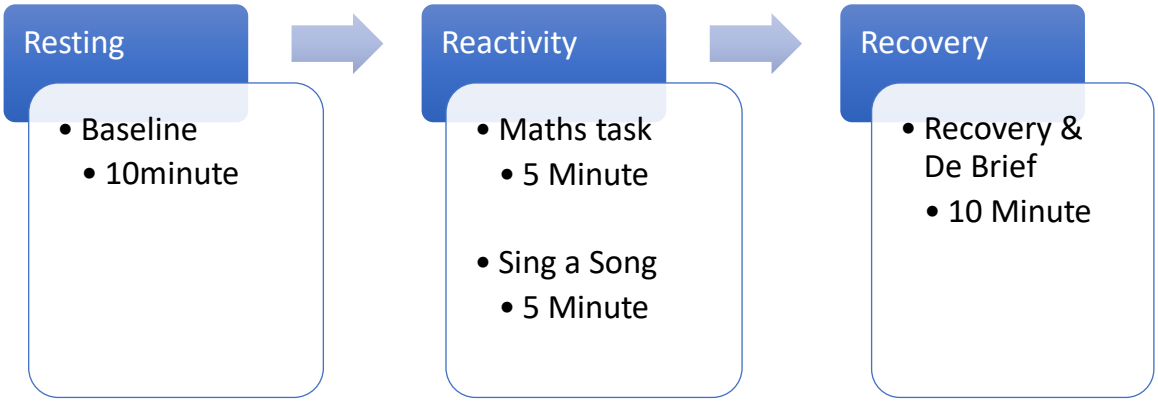
***Figure 3.2**: Sample signal test report*



3.3.9. Laboratory baseline and stressor measurers

In the combined studies an adapted and validated a version of the Sing-a-song stress task (Jump & Dockray., 2020) was used. Firstly, the viability of the sing-a-song task as a technique for inducing mental stress was examined, and then utilized it as a comparative measure to examine the association between laboratory and ambulatory data. Brouwer (Brouwer & Hogervorst, 2014) demonstrated a reliable induction of a stress response with promising avenues using an anticipatory sing-a-song task, while addressing confounds such as standing, speaking, researcher interaction, and also reducing methodological resource load. However, while the SSST has advantages, particularly for short term CV measures (>10 minute epochs), to our knowledge its efficacy utilizing HRV measures, which offer a granulated measure of cardiovascular response, had not yet been demonstrated. The methodological considerations, including adherence to a resting/reactivity/recovery design suitable for HRV analysis are addressed in Study Two. The adjusted the procedures, employing full spectrum ECG analysis, to align with (Laborde et al., 2017) experimental planning, data analysis, and data reporting for HRV type studies using a Resting, Reactivity & Recovery design are detailed.

Figure 3.3: *Overview diagram of SSST timeline and phases*



3.3.10. Baseline

Standardised (Fishel et al., 2007) baseline measurers, including physiological and affective ratings were conducted for all participants to allow within subject comparisons for the reactivity and recovery phases, and also the ambulatory studies. The baseline measurement period consisted of ten minutes of rest sitting in a comfortable chair in the laboratory while wearing the ECG device. Participants are required not to move to reduce artefacts in their baseline heart rate data. Baseline measurement required participants to remain seated on a chair with their knees at a 90-degree angle and both feet flat on the floor. Participants were not asked to keep their eyes closed but were asked not to move if possible. No instructions were given with regards to hand position. Momentary affect was captured using self-report affect scales.

3.3.11. Maths task

The maths task is a computer-based maths task written and presented on PscyhoPy software (Peirce., 2019). Similar to the Trier Social Stress Task (Kirschbaum et al., 1993), participants are required to complete a serial subtraction task while under social evaluative threat. The participants are told that they must achieve a "pass" score to proceed and given a fictional number of iterations of the test that they must achieve. This element is included to increase perceived external social evaluative threat. Scripted comments are added at predetermined times during the task, such as, "Most people have gotten much further by now", to increase the evaluative element. The researcher is present throughout the task viewing the screen and sitting at right angles, in close proximity to the participant, in order to increase immediate perceived social evaluative threat.

A number of experimental elements are managed to maintain homogeneity of presentation during the task; firstly, participants and the researcher sit in the same position and distance apart, secondly, the researcher comments are scripted and carefully managed, thirdly, the researcher maintains a neutral demeanour throughout the presentation. The screen presentation also ensures homogeneity of stimulus as it is central to participant focus throughout the task. Participants are presented with a grey screen with a randomly generated number. Participants are issued with instruction to successively subtract the number 13 and enter their answer in a text field box. The screen also has a countdown timer which starts from 7 seconds. The participant is required to subtract the number and if successful is immediately presented with the resultant number, a reset timer and instruction to continue. In instances where the participant enters an incorrect number, or a timeout of the countdown clock occurs the entire task is reset, and the participant is required to start again. Failure is accompanied by a large bright red text and loud noise instructing them to start again. The measurement period for the task continues for five minutes regardless of the number of failed or successful attempts.

3.3.12. Sing a song stress task

The singing phase of the reactivity section for the included studies is an adaptation (Jump & Dockray., 2020) of the Sing a Song stressor paradigm outlined by Brouwer and Hogervorst (2014) employing anticipation of singing as a primary stressor as a variant of the 'public speaking' task used in the original Trier Social Stress Task (Kirschbaum et al., 1993). The adaptation of the task proposed by Brouwer is used with the inclusion of considerations for HRV measurement.

The outlay for the Singing phase, requires the participant to sit on a chair in front of a microphone and camera. The researcher is present, sitting at right angles, in close proximity to the participant. During the task the researcher maintains a minute-by-minute countdown, increasing to every 15 seconds for the final minute- followed by a ten second countdown at the end of the anticipatory period. Confirmation of the efficacy of this task is presented in Study Two and a detailed account of the findings are presented.

3.3.13. Laboratory conditions.

The laboratory phases of all studies were conducted in the Bio-signal laboratory in the School of Applied Psychology in UCC. The laboratory is located on the ground floor in a windowless, climate-controlled room. Lighting and temperature are controlled at the researchers' discretion and were kept consistent for all participants. At temperature of 17.5 degrees Celsius and a brightness of 400 lumen per m/s were maintained for all participants. The room has two desks and three chairs arranged at right angles. There is also a computer to allow completion of the maths task. The chairs are arranged as close as possible to each other to minimise participant movement, where required, between tasks. A microphone and camera are also permanently fixed in the laboratory.

3.3.14 Momentary Affect Scale – Laboratory.

In order to assess affect during the laboratory collection phases and to match affect to the ambulatory phases a momentary affect scale based on the affective reporting from the Day Reconstruction Method Documentation (DRM) Kahneman, Krueger, Schkade, et al., 2004) was include. For brevity only selected questions were included at laboratory phase and a full list of questions is available in appendix ii

3.3.15. Ambulatory Experience sampling measures - Day reconstruction method (DRM).

The Day Reconstruction Method (Kahneman, Krueger, Schkade, et al., 2004) is used for assessing experiences and affect in daily life. The technique uses retrospective reconstruction using a diary comprised of set time periods to assess participant activity and affect during the previous day. This measure has been included as it represents a non-invasive method of sampling daily experience, and in particular where long term physiological measures are employed it allows uninterrupted sampling without the introduction of experiential confounds.

Although the DRM offers lower possibility of experiential interference, time point accuracy and daily adherence to diary completion can be compromised in DRM type studies (Diener & Tay, 2014). Data captured by the type of experience sampling instrumentation may vary dependant on the specific research question. For example, researchers can include affect scales or measures related to particular behavioural outcomes such as stress events. The DRM has been validated using other experience sampling assessments and has been shown to be a reliable method of for monitoring affect and experience over defined epochs (Dockray, Grant, Stone, Kahneman, Wardle, & Steptoe, 2010).

In the studies presented here the DRM was adapted and hosted online to allow for ease of data entry and monitoring of participation by the researcher. This adaptation was implemented to

address a number of methodological concerns with the DRM, including adherence and participant and researcher burden. The DRM used in the ambulatory study is presented in Study Three and a detailed description of the DRM and the adaptations in the current studies is presented

3.4 Study Procedures

3.4.1 Overview

The following procedure section gives an overview of the procedures for the laboratory for studies two, three, and four included in this PhD. This section includes further detail that may have been restricted in the manuscripts submitted for publication, for example specific procedures regarding technical aspects of heart rate measurement, equipment, and the participant instructions for each phase are included.

3.4.2 Laboratory attendance

Participants were contacted through the email addresses they provided at the initial survey phase of the study and scheduled for laboratory attendance. Participants were informed that the study was examining mental health and heart rate. Participants were also informed of the type of tasks they would be asked to complete in the laboratory and the requirement to wear a HR monitor and fill out a diary for duration of four days for the ambulatory phase. Once fully informed of the scope of the lab and ambulatory studies they were given the option to withdraw without consequence and have all data from the survey phase deleted. Participants then self-allocated to each individual study and were scheduled to attend the laboratory.

Participants were scheduled to attend the laboratory depending on which study they were allocated to. This was determined by work times to capture work/non workdays and also in the employment/unemployed groups depending on employment status. Laboratory sessions were scheduled from 9-1am and 2-5pm. Participants were asked not to smoke and refrain from caffeinated drinks for up to one hour prior to attending the lab. Upon attending the laboratory the participants were asked to fill out the demographic measures and additional self-report measures depending on which study they were allocated to. All measures were completed on a desktop computer, in private, and entered on a google drive survey. Once participants had

completed the self-report measures, they were given a unique identification number to anonymise their data. Data was anonymised to comply with the ethical requirements set out by the School of Applied Psychology Board of Ethics and data protection legislation. Once the surveys were completed participants progressed to the laboratory measurement phase.

3.4.3 Actiheart procedure

After the completion of the self-report measures, the participants were asked to attach the Actiheart Monitors. The Actiheart was attached directly to the skin via two conductive ECG electrodes. The area where the electrodes were attached was prepared to ensure correct adhesion. This required cleaning of the area with sterilised wipes. In cases where hair was present, participants were required to shave a small area to allow the electrode to be attached. If this was required, the participants were given a razor and directed to a private changing facility adjacent to the laboratory. Participants were allowed a short recovery period of fifteen minutes before the electrodes were attached in these cases.

The first electrode was attached at the V2 fourth intercostal space immediately left of the sternum. The second electrode was placed approximately ten centimetres away on the V5 anterior axillary line at the same horizontal plane as the V4. These positions follow the recommendations outlined by the product designer (CamNtech, 2010). Work detailing the placement of patches on signal integrity as set out by Brage et al (2006) for heart rate monitoring in settings where participant movement would be factor was also consulted. The electrode was placed above the pectoral muscle or breast in all cases. The above breast/pectoral muscle position was decided upon to address issues regarding ECG electrode placement and different physical characteristics such as breast tissue or body fat composition (Rautaharju et al., 1998). Although the manufacturer recommends two possible positions (the second is also

aligned with the V2 and V5 but is underneath the breast/pectoral muscle) it was decided to use the single above position in order to maintain consistency of collection method for all participants. Care was taken to ensure the patch placement left the monitor wire taut and extended in the horizontal position to ensure that the monitor movement was minimised and did not interfere with signal integrity.

Once the electrodes were attached the Actiheart monitor was prepared for use. This procedure requires the activation of the unit by placing it on the Actiheart reader and designating a unique user ID and session data for each participant. Session and unit test procedure as set out by the product designers (CamNtech, 2010) were followed to ensure each unit was functioning correctly and had sufficient battery to last the entire data collection period, before use. The ID and session data allow for tracking of each participant's physiological data to the stress task and ambulatory data. Individual participant physical characteristics are inputted to the system at this point also; including participant mass, height, age, gender and estimated max heart rate. Participants are asked to self-report date max heart rate and measures of height, mass and BMI are taken in the laboratory. Upon successful entry of the session data the monitor was attached and a signal test is conducted.

3.4.4 Signal testing

Before commencement of data collection, a period of manufacturer defined signal testing is required to ensure the integrity of the ECG signal. The signal test ensures that accurate recording of the full r-r wave complex of the heartbeat is being captured. This test involves a short recording period to test the integrity of the signal and data being captured for each participant. The test requires the participant to wear the monitor for 12 minutes. The test is comprised of a short period of activity and a rest period. The time period was split in the case

of the current studies into 8 minutes of rest and then a short four-minute walk. This allows for testing of the activity/movement function of the monitor. After twelve minutes the monitor was removed and placed back on the reader.

The Actiheart software provides a specific function for signal quality testing. The software analysis the ECG signal and provides the researcher with an overall pass/fail score for the entire session. In addition, the user is provided with a graphical representation of the entire period highlighted in red or green. Red indicates a period during the test when the signal was inadequate. Green represents an adequate signal. The software provides a function where the user can highlight specific epochs during the test period and examine the signal integrity. The user can then identify if signal loss is due to movement or adhesive patch placement. The user can then make modifications to address these issues if necessary, either by moving the adhesive patches or ensuring the monitor is correctly attached. Throughout the various studies there were instances where these adjustments were required. In these cases where the signal test was not passed, placement adjustments were made to the patches and the participants were asked to complete the signal test again. A cut off point of three tests was pre-determined prior to commencement of the studies to reduce participant burden, however all participants completed successful signal tests before the second signal test. Once the signal test was successfully completed participants were allowed to progress on to the laboratory and ambulatory study. Participants were provided with additional patches and instruction information on the Actiheart use.

3.4.5. Long-term recording

Once the signal test was completed the Actiheart was reattached to the participants. The Actiheart monitor remained on the participant for the entirety of the study period from then on.

This included the laboratory task and (or) the four-day ambulatory phase. To facilitate this a long-term recording session was loaded onto the Actiheart monitor (Details of the different types of recording session allowed by the Actiheart monitor; Short term and long term, and the implications for recordable data are detailed in the materials and measures section and the data analysis section of the current methodology). User ID and session ID were logged. A 15 second epoch for beat analysis and activity was used. Once the set-up procedure was complete the participants were progressed on to next phase of each study

3.4.6. Baseline procedure

Upon successful completion of the signal testing all participants carried out baseline measures. Baseline measures were conducted with participants in a seated position with their legs at 90 degrees to their body. Participants were instructed not to move for 10 minutes while the baseline procedure was completed. Movement was monitored via accelerometer data in the Actiheart monitor. All participants successfully completed a baseline measure before continuing on to the reactivity (stressor) phase.

3.4.7. Maths task procedure

Prior to beginning the maths task participants are given instructions on its completion. Participants were told they would be required to pass an "intelligence" test in order to participate any further. Participants were then told they had to maintain an 80% standard to be successful. The importance of completing the task and possible consequences, such as not being able to proceed and significance of the maintaining successful scores were emphasised by the researcher. The test conditions for the maths task require the participants to sit in front of a computer screen and complete a serial the subtraction task. The researcher sat on a chair directly next to the participant to increase the social evaluative pressure. The maths task was

completed for all participants as described in the materials and measures sections for Studies Two and Four

3.4.8. Anticipatory Sing a song procedure

Immediately following the maths task, the participants were informed by the researcher that they would be required to complete an additional task as part of their "evaluation" This test would require them to 'sing the song' into a microphone in front of the researcher in the same seated position as the maths task so as to eliminate movement. All participant for studies two and four completed the anticipatory singing task as described for each study respectively.

3.4.9. Recovery phase

All participants for studies two and four completed a recovery phase. The recovery phase matches the procedure for the baseline measures. The participants are required to sit in the same position and a ten-minute recovery period is measured.

3.4.10. Laboratory De-brief

Once laboratory procedures were completed the participants were de-briefed. Participants were informed that the task was not an intelligence test, but a maths task deliberately designed to elicit a stress response. They were also informed that their progress was not contingent on passing the task and that the random average scoring did not reflect their ability in any way. Care was taken to ensure that any residual stress had passed and that participants fully understood the nature of the study and why the initial deception was required. For the participants who were only taking part in the laboratory phase this marked the end of their participation. The participants that were assigned to the ambulatory study moved on to the ambulatory phase.

3.4.11. Ambulatory phase procedure

The ambulatory data collection period consisted of four full consecutive 24-hour periods for each participant. The commencement of the ambulatory study period was determined by each participant's reported work/non-workdays. The common feature was a four-day collection period spanning two workdays and two rest days for each participant. Owing to the varying employment demands this meant the days of the week differed for many participants. Data collected during the ambulatory phase included the HR data and the data from the DRM. Actiheart data was collected on a long-term recording session as described for Studies Two and Four respectively. Further detail on how the physiological data were selected and analysed is presented in each study with further detail in the data analysis section in this chapter. The adapted version of the day reconstruction method (Kahneman, Krueger, Schwarz, et al., 2004) as described is presented in detail in Study Three and was used to collect the experiential data. Further detail on the ambulatory procedure is presented in the following section.

3.4.12. Participant pre ambulatory briefing.

Participation in the ambulatory study required detailed information and instructions to be given to the participants. This included information on how to successfully complete the Day reconstruction measures and relevant information on Actiheart. The information was provided to the participants in the form of information packs (appendix i). Time was taken, prior to participants leaving the laboratory to brief participants on the contents of the pack and the requirements for completing the DRM measures.

3.4.13. Actiheart device information

Participants were provided with information about the Actiheart monitor. This information was added after the pilot phase to participant feedback asked such as "could the device record

audio" and "what do I do if the device falls off". Information was provided that described the exact recording capabilities (movement and HR) of the Actiheart device and instructions on how to deal with mishaps in electrode attachment including instructions on how to remove the Actiheart monitor and electrodes at any time if they wished to do so. Participants were also informed that they could remove the monitor and stop participation at any time if they wanted to do so. In these instances, they were asked to record the time in the DRM online.

3.4.14 Day Reconstruction Method information.

As the day reconstruction method required a significant amount of effort on the part of the participant (See DRM procedure below) time was taken in the laboratory before the participants left to describe the procedure and provide information. The DRM included a daily diary, a website survey log on instructions and packets one and four to be completed before leaving the lab. Following the pilot study, it was decided that participants should complete a practice daily diary before they left the lab. Once the participant and researcher were satisfied that they understood the procedure they progressed on to the ambulatory phase.

3.4.15. Participant briefing

Participants in studies three and four were require to carry out the Day Reconstruction Method demographic and daily measures. Participants were provided with the briefing information and asked to complete a practice version of the day reconstruction method in the laboratory. For the studies presented here the DRM was separated by packet. Packets one and part of packet four were completed as part of the initial survey when the participants attended the laboratory. Packets two and three along with the daily questions from packet four were completed retrospectively, concerning the day before, at the end of each study day.

3.4.16. Demographic Information

Packet one and sections of packet four in the DRM represent demographic and employment information. These sections were included in the self-report survey given to participants in the lab. These data were later collated with that for DRM data for each participant.

3.4.17. Daily diary procedure

The participants for studies three & four were required to complete the daily diary at the end of each day during the study period. Participants were instructed to set aside time and sit in a quiet room alone and fill out each section. Packet two (handwritten personal diary) of the DRM is the basis for the data the participants provide for packet three (the web form) of the day reconstruction tool. Packet two comprises of a daily diary, seen only by the participant, that facilitates the retrieval of autobiographical information. Participants were informed that the diary section is completely confidential. Kahneman (2004) stresses the methodological importance of both completing the diary section prior to progressing to packet three and its confidential nature to increase accurate recall but also reduce omission of information. Diary information for the day of completion represents entries for the twenty-four-hour period immediately before.

Participants were instructed to set aside time after completion of the daily diary to complete the DRM online. Participants were instructed to access the system in a quiet room and to have their diary information readily available. Participants then logged on to and completed the DRM measures online, taking care to detail each the required information and affect scales for each episode. Participants completed the DRM for the four study days and once participation was complete the data was collated with the Actiheart and survey data and subject to data analysis. Full data analysis and treatment of the data is outlined in the data analysis section.

3.5. Data analysis

For studies two, three, and four the data analysis is described in each case respectively. This section provides further detail to contextualise the treatment of the data.

3.5.1. ECG Artefact and error correction

ECG data is captured and sampled form the Actiheart device as described in the materials and measures section. The Actiheart software includes a proprietary algorithm that cleans and corrects HRV data for artefacts prior to presenting it to the users. As HRV data is subject to influence from myriad confounding factors, for example, movement, disruption of electrode attachment, and missing or misattributed beat detection (Kligfield et al., 2007). These errors can be introduced at various stages of detection, either by the device itself or the software that is being used to interpret the data. Studies examining HRV therefore need to account for these errors and a number of methodological solutions exist to account for these confounds. Firstly, using detailed manual examination of HRV signals for each participant for each measurement period using software such as Kubios™ (Tarvainen et al., 2014). This approach gives the researcher the most granularity and control over signal detection and allows precise examination of all artefacts. However, this approach carries a considerable resource load with manual examination of each participant's HR data. In this case individual missed beats and artefacts are examined, for example with movement and respiratory data, and decisions about inclusion/exclusion are made. Secondly, other software such as ARTiiFACT™ (Kaufmann et al., 2011) can be used with data sets to correct for errors. These solutions include automated corrections for missed or ectopic beats, such as interpolation of missing data from the mean/median distribution of related scores for each participant, and have been demonstrated to reliably shown to account for missing and erroneous data (Lipponen & Tarvainen, 2019). The use of software such as this allows for considerable time and labour savings on behalf of

the researcher, however they do require that the artefact correction is trusted to the software and less fine-grained level of control from the previous method. Finally, artefact management can be integrated and supplied with the device software itself, and many manufacturers have integrated these solutions with their hardware. These solutions offer the most methodologically expedient solutions for researchers, with considerable savings in workload, however as the quality of these devices varies considerably not all are suitable for use in research or clinical settings. In addition, the usual issues regarding proprietary software exist, for example there may not be transparency into the background processing of data and no opportunity for a granular analysis of the data if required. Therefore, platforms with hardware that permits the extraction of full spectrum ECG signal and offers integrated software solution for data sorting and error correction are desirable.

For the included studies here the data error correction and sorting were carried out using the Actiheart device proprietary software and then exported so SPSS statistical software. The device captures full spectrum ECG that offers its own user-controlled artefact correction procedure. In addition, the software has an “auto clean” function that carries out cleaning and sorting automatically. Where this is selected the data are sorted and presented to the user with a “quality index” of 0-1. Although this index provides a scale that can be used to parse data by quality of signal, when considering the use of Actiheart for the studies here, what the levels of “quality’ indicated in relation to Artefact correction was investigated it was determined that no included epoch, either in the laboratory of ambulatory context, would be included unless a perfect score for quality was indicated. Although this led to the exclusion of some epochs from the HRV analysis, this more conservative approach was favoured to ensure the integrity of the data.

3.5.2. Actiheart accelerometer data and movement

Movement is captured by the Actiheart and the device has been validated in relation to other devices for movement and energy expenditure calculations that include HR data (Barreira et al., 2009). Examination of movement data is permitted using indices returned by the ECG measurement device and presented with ECG raw data as an "Activity scale count". Although study of HR and in particular ambulatory contexts can include the use movement data and extrapolate its impact on HR, for example in long-term 24hr capture of HRV it can be difficult to determine exactly how activity is related to specific parameters for movement. In addition, different device platforms may apply scales to SI units, such as metres per seconds (m/s) while others, such as Actiheart, use undefined units of movement. Given the breadth of devices available the range of movement indices may not be comparable, moreover it is difficult to clearly define and standardise "movement". Therefore, for the purposes of both the laboratory and ambulatory studies an additional quality control procedure was implemented at data analysis phase and only epochs that showed no movement were included in the HRV analysis. In the laboratory context this was employed across the entire procedure, and although the researcher was present to observe the participants, the accelerometer data provided an additional level of control for data inclusion. For the data captured in ambulatory contexts each epoch included in the analysis was required to have no movement. In addition to eliminate confounds related to movement prior to each ambulatory epoch, the five minutes prior to the included epoch was also required to have no movement. The constraints on movement data were applied to allow for the proposed comparison of the laboratory and ambulatory contexts.

3.5.3. Study four - Ambulatory affect and HRV sampling

For inclusion in the final data set the ambulatory data was subjected to the following set of predetermined criteria:

1. The ambulatory measurement period for each participant was comprised of 4 x 24hr segments. This was grouped in to 288 x 5-minute segments, with 4 days per participant, resulting in a typical sample of 1152x5 minute epochs per participant (Owing to battery capacity, that is related to and influenced by the amount of activity and number of heartbeats per day, the constraints of the Actiheart device, and given the required settings for the current study, a number of participants only returned 3x day Actiheart data).

2. Affective data as reported by the participants in the DRM were used to assign participants' episodes to either a positive affect or a negative affect condition. Initially the study aimed to include six negative and six positive samples for each participant, however upon examination of the data it was determined that participants did not report enough negative samples to complete the data set. Therefore, only positive conditions were included.

3. For each participant the remaining positive affect episodes were entered into a random number generator (Google random number generator) and this was used to randomly sample across the identified measurement period.

4. After identification of suitable epochs for each participant the social conditions and location were examined. Participants were required to be alone but could be a at number of locations (for full list of locations included in DRM). For each episode participants were asked to self-report qualitative data also to account for any confound, for example, data were excluded because participants were watching sporting events, or similarly evocative programmes on TV and the influence of this would be difficult to determine.

5. Once the epochs passed the criteria for affect and location, the accelerometer data for each epoch was exported examined. A threshold of NO movement (for both the included epoch and

the one preceding it) and a 100% quality rating for HRV signal was required for the epoch to be included.

6. Once the randomised data satisfied all of these criteria it was included in the multilevel model for analysis.

Table 3.2: *A sample of checklist for context for positive affect ambulatory HRV*

Criteria	**Report**
Participant number	23
Day number	3
Randomly identified episode number	166
Time of day	19:45
Episode -social interaction	Alone
Episode-location	Home
Accelerometer data	No movement
Qualitative data	No confounds indicated
Heart rate data quality	Above threshold
Heart rate data	Segment 3
Randomly chosen 5 min segment	No interruptions

3.6 Pilot phase

3.6.1. Pilot phase overview

A pilot phase of Studies Two, Three, and Four, was conducted to test the feasibility of the study design and also to allow the researcher to familiarise themselves with the procedure and the various measurement devices and software employed in the studies. This included, the Actiheart set up and electrode placement, Actiheart signal testing, and downloading and processing of data retrieved from both phases. In addition, the feasibility of the four-day measurement period was tested for both physiological measures and self-report measure. Details of each phase and the implications for the final study designs are outlined.

3.6.2. Pilot study Participants

N=4, Male= 2, Female = 2, were recruited from the general population using convenience sampling following the procedure previously outlined.

3.6.3. Short interview of participant experience

Once the pilot phase was completed the participants were de-briefed and interviews were conducted to assess their experience of the study design. Their responses were integrated into the final study design presented in Studies Two, Three, and Four, and this resulted in a number of adjustments to the overall methodologies.

3.6.4. Pilot findings – Laboratory phase

The pilot phase of the study demonstrated the potential initial efficacy of the laboratory procedure to elicit a stress response from the participants. Reactivity was demonstrated in a resting/reactivity/recovery profile as expected with HRV data indicting initial changes in HRV across tasks. Resultant from the examination of the heart rate data and the participant

interviews during the pilot phase of this study a number of changes were implemented for the main phase of the laboratory study, including, a longer lead in time before baseline measures were taken to allow the participants to acclimatise to the laboratory environment and the researcher. Self-report momentary affect scales were included to monitor participants perceived affect during the tasks. A change in seating position between the respective tasks was changed to the participants' sitting in the same chair so as to eliminate movement between task positions. Finally, a swivel chair that had been used was removed due to participants movement that introduced HR confounds.

3.6.5. Pilot findings-Ambulatory phase

The pilot study tested the designs for studies three and four across a four-day (24 hr) measurement period. The pilot examined participant experience, measures of heart rate, and the daily diary (DRM), in context, to determine the feasibility of the study protocols. The pilot phase also served to help the researcher to determine the logistical implications for the day reconstruction method, both for the researcher and the participants, and to make any required adjustments to the protocol for the full-scale study.

Resultant from the pilot phase a number of amendments were made to the study design including, amendments to DRM instructions, where more time was allotted to explaining the DRM to participants, and the inclusion of a trial run of the DRM in the presence of the researcher. Finally, additional ECG patches along with training for participants how to reattach them were included. The pilot phase coincided with a very warm period of weather that resulted in ECG patches falling off due to perspiration. Subsequently procedures were amended to capture these occurrences in the DRM and account for them at data analysis.

3.7 Ethics

The studies included here were submitted for ethical review by the Ethics Committee in the School of Applied Psychology in University College Cork. All studies were granted ethical approval.

Chapter 4: Study two - Cardiovascular responses to stress utilising anticipatory singing tasks.

4.1 Abstract

Models of psychobiological stress reactivity have a foundation in the measurement of responses to standardised stress tasks. Tasks with anticipatory phases have been proposed as an effective method of stress induction, either as a stand-alone task or replacement constituent elements for existing stressor paradigms. Tasks utilising singing as a primary stressor have been proposed but the efficacy of these tasks have not been demonstrated while maintaining adherence to a resting/ reactivity/recovery framework desirable for HRV measurement. This study examines the viability of an anticipatory sing-a-song task as a method for inducing mental stress and examines the utility of the task with specific reference to measures of cardiovascular reactivity and recovery activity, and standard protocols to examine HRV reactivity and recovery. Participants completed a dual task with a maths task and an anticipation of singing component. Responses were examined according to a resting/reactivity/recovery paradigm and the findings indicate that the sing-a-song stimulus is effective in generating a stress response. Significant differences in heart rate and self-reported stress between baseline and stressor conditions were detected, with greater magnitude differences between baseline and anticipatory phases. This study has demonstrated the viability of the anticipation of singing as a standardised stressor using cardiovascular measures and has described variants of this task that may be used for repeated measures study designs.

Outputs

Published as: Jump, O., & Dockray, S. (2020) Cardiovascular responses to stress utilizing anticipatory singing tasks. Journal of Psychophysiology (Preprint) https://doi.org/10.1027/0269-8803/a000269

Jump, O., & Dockray, S (2017) A variant of the Trier Social Stress Task using Sing-A-Song and computer-based tasks Psychological Society of Ireland, Division of Health Psychology Annual Conference.

Presented as: Jump, O., & Dockray, S (2017) Social interactions and heart rate variability: From the lab to the living room: Psychological Society of Ireland, Division of Health Psychology Annual Conference.

Contributions

The lead researcher and author on this study is Owen Jump. The work was supported and reviewed by Samantha Dockray.

4.2 Introduction

Models of psychobiological stress reactivity have a foundation in the measurement of responses to standardised stress tasks. To delineate these responses, a dedicated literature has examined the efficacy and methodological applications of stress tasks, with a focus on acute reactivity via chronic psychosocial stress (Chida & Hamer, 2008). These stressors have included cognitive tasks such as, serial subtraction, public speaking, stress interview, emotion induction and social evaluative stress. These tasks have been utilised to elucidate differential responses in biology, such as hypothalamic pituitary adrenal (HPA) activation, via cortisol, ACTH, vasopressin, DHEA, and cardiovascular activation (Allen, Kennedy, Cryan, Dinan, & Clarke, 2014). Demographic and psychosocial variables have also been examined in relation to stress task efficacy including, age and gender (e.g. Kudielka, Buske-Kirschbaum, Hellhammer, & Kirschbaum, 2004) and pubertal stage, (e.g. Sumter, Bokhorst, Miers, Van Pelt, & Westenberg, 2010). Further study has sought to examine pathophysiological processes of physical illness brought about by sustained hypertensions in response to stress challenge (e.g., Thayer, Yamamoto, & Brosschot, 2010) and psychopathologies, for example, depression (Burke, Davis, Otte, & Mohr, 2005; Young, Lopez, Murphy-Weinberg, Watson, & Akil, 2000) and anxiety (Gerra et al., 2000; Young, Abelson, & Cameron, 2004).

In order to better model ecologically valid stressors, tasks that utilise a combination of psychosocial procedures such as the Trier Social Stress Task (TSST) (Kirschbaum, Pirke, & Hellhammer, 1993) have been developed. The TSST has emerged as the most widely used stress test (Kudielka, Hellhammer, & Kirschbaum, 2007), and has been used to reliably induce stress, enabling the relationship between stress and physiological responses to be examined (Allen et al., 2017). Although widely employed by researchers, a number of challenges of the TSST exist, including, methodological resources load (Allen et al., 2017), potential confounds

such as movement (standing) and speaking (Brouwer & Hogervorst, 2014), and habituation challenges, when participants are required to complete repeated measures, in particular in relation to HPA activation (Schommer, Hellhammer, & Kirschbaum, 2003; Arvidson, Sjörs, & Jonsdottir, 2017). Reducing habituation effects presents a need to identify and validate for replacement task elements in existing stressor paradigms, such as variation of maths task or novel types of social evaluative threat, although it is noted that while habituation has been observed for HPA activation, CV reactivity and blood pressure appear less susceptible to habituation effects (Boyle et al., 2016). Given the challenges indicated above, there is a need to identify and verify possible replacement elements or stand-alone tasks with demonstrated equivalence of efficacy.

Brouwer and Hogervorst (2014) proposed the Sing-A-Song Stress Test (SSST) as a task for inducing mental stress with the aim of addressing a number of the challenges inherent in the TSST such as movement, standing and speaking. The SSST elaborates on a methodology proposed by Hoffman (Hofmann, Moscovitch, & Kim, 2006) designed to examine reactivity in shy/non-shy participants using anticipation of a singing task to induce a stress response. Hoffmans' original methodology informed participants that they would be required to give a speech and to sing-a-song. The SSST further developed this task using computer-based stimuli while standardizing presentation and the inclusion of time locking of stimuli to reduce confounds. The task was efficacious, but the timed periods did not allow examination of HRV measures. This study proposes the examination of the viability of the anticipatory singing task protocol in combination with a subtraction maths task order to adhere to standardized methods for measurement of cardiovascular activity, including a Resting/Reactivity/Recovery (RRR) structure (Laborde, Mosley, & Thayer, 2017; Shaffer & Ginsberg, 2017), which allows for delineation between tonic and phasic HRV between conditions. This design has been proposed

by Laborde as a standardised procedural outlay to allow for examination of change in HRV across conditions.

Cardiovascular activity is a widely used measure in psychophysiological science, and changes in cardiovascular parameters are related to individual differences, such as personality (Huang et al., 2013) and risk factors for health and morbidity (Kemp & Quintana, 2013), and are relatively easily measured in comparison to other biological systems (e.g., the cortisol response). The most commonly used indicators of cardiovascular activity are blood pressure, heart rate reactivity and heart rate variability. The reactivity hypothesis (Blascovich & Katkin, 1993) posits that heightened cardiovascular reactivity, typically reported as increase of beats per minute (BPM) and blood pressure (mm HG), observed in response to stress can promote sustained hypertensions that contribute to the development of cardiovascular disease (Thayer, Åhs, Fredrikson, Sollers, & Wager, 2012; Allen et al., 2017; Everson et al.,1997). Heart rate variability (HRV) has also been widely used in psychophysiological study (Laborde, Mosley, & Thayer, 2017; Quintana & Heathers, 2014). HRV describes a range of measures, including time domain and frequency components of the heartbeat that have been associated with affect (Kreibig, 2010; Porges, 2001) and social functioning (Porges, 2001; Quintana, Guastella, Outhred, Hickie, & Kemp, 2012). Typically, time domain measures describe variability in successive R –R peak intervals or variations in Inter-Beat Interval (IBI). Root Mean Square of Successive Differences (RMSSD) reflects the beat-to-beat variance in IBI for cardiovascular activity and has been associated with HF power (Shaffer & Ginsberg, 2017). In addition, frequency domain measures are based on the assumption that each range is indicative of the 'power' or signal energy for that given band; Ultra Low Frequency (ULF), Very Low Frequency (VLF), Low Frequency (LF) and High Frequency (HF). ULF has been associated with circadian rhythm and metabolic demands (Shaffer, McCraty, & Zerr, 2014). VLF has been linked to physical movement (Bernardi, Valle, Coco, Calciati, & Sleight, 1996) and

parasympathetic activation (Taylor, Carr, Myers, & Eckberg, 1998). The LF band has been associated with baroreceptor activity, and subsequently blood pressure (Shaffer & Ginsberg, 2017), and has been proposed as an index of sympathetic outflow (Reyes del Paso, Langewitz, Mulder, van Roon, & Duschek, 2013), although inclusion of LF is questionable and disagreement exists regarding its use as an index of sympathetic activation (Billman, 2011, 2013; Heathers & Goodwin, 2017). HF components of HRV have been reliably associated with vagal tone (Porges, 2007), and is therefore correlated with stress, permitting its' use as measure to reflect parasympathetic activity (Laborde et al., 2017). Finally, low frequency/high frequency ratio has been proposed as an index of autonomic balance (ChuDuc, NguyenPhan, & NguyenViet, 2013; Quintana & Heathers, 2014), although as mentioned, because the inclusion of LF bands is questionable, employing the LF/HF ratio carries related concerns. While debate exists regarding the various HRV components, for the purposes of the current study time domain measures of BPM, IBI, RMSSD and frequency components of HF are including following recommendations by Laborde (Laborde et al., 2017) and Quintana (D. S. Quintana et al., 2016) for experimental planning and data reporting using HRV measures

In this study, we examine the inclusion of an anticipatory sing-a-song task as a task element along with a subtraction maths task for inducing mental stress while retaining adherence to a resting/reactivity/recovery design suitable for HRV analysis. Initial findings from the SSST, as demonstrated by (Brouwer & Hogervorst, 2014) indicate a reliable induction of a stress response with promising avenues using an anticipatory sing-a-song task that addresses confounds such as standing, speaking, researcher interaction, and also reduces methodological resource load. While the SSST has advantages, particularly for short term CV measures (<10 minute epochs), to our knowledge it has not been examined using more granulated measures of HRV that offer advantages in interpretation of autonomic function (Quintana & Heathers,

2014). A number of methodological considerations arise when considering the use of the SSST with HRV measures, such as matched timing of epochs of sufficient duration for HRV measures. This study employs full spectrum ECG analysis and presents reactivity measures (BPM), time domain measures - beat to beat interval (IBI) and Root Mean Square of Successive Differences (RMSSD), and frequency domain measures of high frequency (HF). A number of adjustments are suggested to align the procedures of the SSST with recommendations (Laborde et al., 2017) for experimental planning, data analysis, and data reporting for HRV type studies utilising a Resting, Reactivity & Recovery design with anticipation of singing stressor stimuli included along with established stressor protocols.

4.3 Method

4.3.1 Participants

Participants (n=45, Male= 21, Female = 24, aged 19 – 65, Mean 31.89 SD=11.34) were recruited from the general population using convenience sampling. Calls for participation were advertised using social media and university email. Participants were contacted after initial expression of interest and given a schedule for lab attendance.

4.3.2 Design

This study uses a within-subjects laboratory-based design to evaluate the viability of an anticipatory singing task utilising cardiovascular measures. The task includes two components, a standardised math task followed by an anticipatory 'sing-a-song' task in Resting / Reactivity/ Recovery segments. Patterns of HRV are used to assess the efficacy of the two task components, separately, in combination, and compared to baseline and recovery.

4.3.3 Materials and measures

Physical environment and laboratory conditions.

Measurements were taken in the laboratory, located on the ground floor of the psychology building. The room has no windows except a one-way observation mirror on one wall. Lighting and room temperature were kept consistent for all participants and was managed at 18.5°C. The room is equipped with visible camera and microphone recording equipment. Participants sat on a stationary office chair for the measurement period to minimise movement.

Physical measures & Confounds

Physical characteristics of the participants were recorded in the laboratory; Participant height(cm) and weight (kg) were measured, and BMI was calculated. Age & gender were self-

reported. Participants self-reported physiological and lifestyle factors (both stable and transient) that could cause potential confounds in the heart rate data medical conditions or medications including, cardio-active medications, antidepressants or anti-hypertensives. These measures were used to preclude participants from the study.

ECG measurement

ECG measurement for this study was carried out using the Actiheart™ (Brage, Brage, Franks, Ekelund, & Wareham, 2005) measurement device. Actiheart is a validated (Kristiansen et al., 2011) non-invasive method of measuring physical activity and energy expenditure (PAEE) (Crouter, Churilla, & Bassett, 2008). The Actiheart is attached over the skin, via an electro dermal patch, and is used to measure ECG heart rate and movement. The device is capable of capturing the full QRS spectrum for defined epochs and uses the resultant data in its' proprietary software to calculate and resolve HRV measures using the Pan Tompkins QRS detection algorithm (Pan & Tompkins, 1985). R-R intervals for at least 16 heart beats are captured for each epoch and averaged. Artifact correction is carried out by deletion by the of values outside of a +/- 25% range and removed and the remaining data re-averaged. In addition to assessment and correction of the r-r signal quality the software presents an additional quality index (range 0-1) for each epoch that can be examined by the user and excluded if necessary. For the purposes of the current study this additional quality control was included and any epochs with a sub optimal quality rating (Less than 1/ perfect score) was excluded from the analysis. The Actiheart device also returns measures of physical activity for the participant. The device measures movement in 'counts of unit per time'. This time/second measurement is then applied by the proletary software to the user defined epoch (CamNtech, 2010; 2018) returning the activity measure for the epoch. Only epochs where no physical activity was detected are included.

Resting - Baseline

Baseline measures (Fishel, Muth, & Hoover, 2007) were collected once the participant was fitted with the ECG device and sitting resting in a chair for 10 minutes. Participants are required not to move to reduce artifacts in their baseline heart rate data. Participants were required to remain seated on a chair with legs their knees at a 90-degree angle and both feet flat on the floor. Participants were not asked to keep their eyes closed but were asked not to move if possible. No instructions were given with regards to hand position. Measures of HR and movement were recorded during the baseline period.

Reactivity – Maths task

The maths task component is a computer-based maths task written and presented on PsycoPy Software. Similar to the maths element of the standard Trier Social Stress Task (Kirschbaum et al., 1993), participants are required to complete a serial subtraction task while under social evaluative threat provided by constant observation of a researcher, seated close to the participant. The maths task is presented to participants and they are told that they must achieve a fictitious score and the number of iterations of the test that they must achieve. This element is included to increase perceived external social evaluative threat. Scripted comments, such as "Most people have done much better by now" were added at predetermined times during the task to increase the evaluative element

Reactivity- Singing task

The singing phase of the reactivity section is an adaptation of the stressor paradigm outlined by Brouwer & Hogervorst (2014) employing 'singing' as a primary stressor as a variant of the 'talking' task used in the original Trier Social Stress Task (Kirschbaum et al., 1993). The variant of the Sing-A-Song used in the current studies represents an adaptation of the task

proposed by Brouwer with the inclusion of considerations for heart rate variability measurement. In the current study the variation introduces an anticipatory stress phase induced by a surprise instruction to sing a song. The singing phase commences once the researcher hands the participant a copy of song lyrics. It is indicated to the participant that they will be required to sing the song to the researcher following a 5-minute preparation period. The singing-phase, requires the participant to sit in the same position as the baseline and maths phase, with the addition of a camera and microphone, introduced by the researcher to "record the singing session" . The researcher is present, sitting at right angles, in close proximity to the participant as they prepare to sing. A verbal minute by minute countdown is included – increasing to every fifteen seconds for the final minute- followed by a ten second countdown at the end of the anticipatory period. At the end of the anticipatory phase the participant is told that they are not required to sing. The five-minute preparation period constitutes the entire singing anticipatory phase.

Recovery phase

A five-minute recovery phase was completed by all participants. The recovery phase, which mirrored the baseline measurement was included to allow a secondary comparison to the task situation (Quintana & Heathers, 2014). Recovery consisted of five minutes of rest sitting in a chair in the laboratory while continuing to wear the ECG device. Participants are again instructed not to move to reduce artifacts in the HR data.

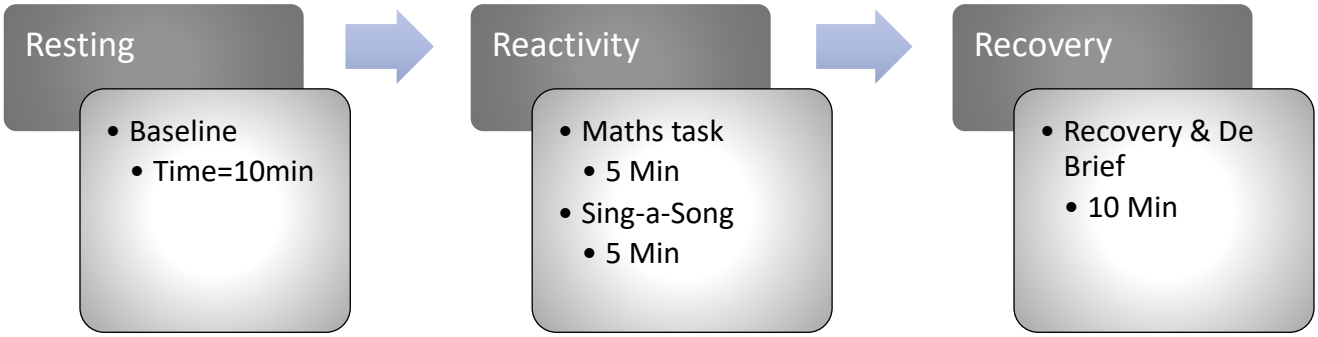
Self-report affect

A self-report affect scale was included at all periods of measurement during the laboratory task. The affect scale requires participants to self-report five dimensions of subjective affect on a rating of 1-5, including, measures of relaxation, stress, control and anxiety the self-report

measures were included to allow for quantification of participant self-perceived affect during the task.

Figure 4.1: *Diagram of SSST Timeline and phases*

4.3.4 Procedure



Laboratory attendance

Participants were contacted and asked to attend the laboratory. Participants were required to refrain from caffeinated drinks and smoking for at least one hour prior to attendance. Participants were greeted by a researcher dressed in a white laboratory coat in a friendly but neutral manner. Participants were then escorted to the laboratory. Once the consent procedures were completed, the participants were asked to complete the self-report measures. Data was anonymised to comply with the requirements set out by the of Ethics committee and data protection legislation.

ECG procedure

Actiheart is attached directly to the skin via two conductive ECG electrodes. The first electrode was attached at the V2 fourth intercostal space immediately left of the sternum. The second electrode was placed approximately ten centimetres away on the V5 anterior axillary line at the same horizontal plane as the V4. These positions follow the recommendations outlined by the product designer (CamNtech, 2010) and by Brage et al (Brage et al., 2006) for patch location

on signal integrity. Although the manufacturer recommends two possible positions (the second is also aligned with the V2 and V5 but is underneath the breast/pectoral muscle) it was decided to use the single above position in order to maintain consistency of collection method for all participants due to different physical characteristics such as breast tissue or body fat composition (Rautaharju, Park, Rautaharju, & Crow, 1998). Once patch placement was completed the participants moved on to signal testing.

Signal testing

A signal test is required to ensure the integrity of the signal and that accurate recording of the full r-r wave complex of the heartbeat is being captured. This involves a short recording period of 12 minutes for each participant. The Actiheart software provides a specific function for signal quality testing. The software analysis tests the signal and provides the user with an overall pass/fail score for the entire session. All participants successfully completed sufficient signal quality tests.

Baseline

Baseline measurements were completed by all participants (5 minutes acclimatisation / 5 minutes measurement) of rest sitting in a comfortable chair in the laboratory prior to engaging in the stressor phases

Maths task

Participants were informed they would be required to pass an “intelligence” test in order to participate any further. Participants were then told they had to maintain an 80% standard to be successful. The importance of completing the task, possible consequences of failure and significance of the maintaining successful scores was emphasised by the researcher. The maths

task procedure takes 5 minutes to complete. A number of experimental elements are managed to maintain homogeneity of presentation during the task; firstly, the researcher sits in close proximity to increase social evaluative elements of the task, secondly, the researcher comments are scripted and carefully managed, thirdly, the researcher maintains a neutral manner throughout the presentation. The screen presentation ensures homogeneity of stimulus as it is central to participant focus throughout the task.

Numbers are presented to the participants with a blank box to type answers. In order to increase the task difficulty, each task is timed, and timing is represented by a clock countdown on the screen. If participants are slower than the allocated time it represents a failure of that particular task. Each failure/pass is also accompanied by a sound and image, with a low sound and green tick/correct mark for a pass, and a loud noise and visual red X for a fail.

Participants progress was observed by the researcher and scripted comments were mentioned at random times, regardless of participant progress. These included statements such as, "As you can see, most people have a slightly higher score by now" and "We need to speed up your progress if you are going to be able to proceed in the experiment". Participants were also reminded by the researcher throughout the task of a fictitious average score of other participants. All participants completed the maths task.

Sing a song task

Immediately following the maths task, the participants were informed by the researcher that they would be required to complete an additional task as part of their "evaluation". This test would require them to 'sing the song' into a microphone in front of the researcher. Participants were also told that a camera and microphone would be used to record them singing and this would be later used as part of their evaluation. Participants were then given a copy of the song they would be required to sing and instructed to familiarise themselves during a five-minute

waiting period. During this period, the researcher was in close proximity and the countdown provided. When the countdown period ended participants were informed they would not have to sing. Participants then moved to the recovery phase.

Recovery phase

A recovery phase of five minutes followed directly after the stressor periods. During this time, the participants were required to sit still to mirror the baseline measurement. Participants were debriefed in line with ethical procedures for laboratory stress protocols and fully informed of the nature of the study.

4.3.5. Data Analysis

HRV data for this study was accessed using the Actiheart proprietary software, the software enables researchers to visualise and export HRV data. The software uses returns measures of beats per minute (BPM), Inter-beat interval IBI (Maximum, minimum and average measures of IBI are returned), root mean of successive square differences (RMSSD), standard deviation of NN intervals (SDNN). Also, frequency domain measures of High Frequency (HF), Very Low Frequency (VLF), Low frequency (LF) and finally Low frequency/High frequency ratio (LF/HF) are captured. For the purposes of the current study, measures of BPM, IBI, RMSSD and HF components only are reported, this is owing to disagreement and reliability regarding the LF components. Actiheart returns ECG data for defined epochs and its' proprietary algorithm sorts and cleans data prior to presenting it to the user. Examination of movement is permitted using movement indices returned by the ECG measurement device. For the purposes of the current study an additional quality control procedure was implemented at data analysis phase and epochs with sub optimal quality ratings was excluded from the analysis.

4.3.6 Ethics

This study was approved by the local University Ethics Committee.

4.4 Findings

4.4.1 Overview

The following results describe the mean differences between test conditions for time domain and frequency components of HR. Repeated measures ANOVA's were conducted to compare of BPM, IBI, RMSSD and High Frequency HRV during baseline, maths stress task, Sing-a-Song anticipation, and recovery phases. Results for each measure are detailed including self-report momentary affect scales. Gender and age comparisons are also included.

Table 4.1. *Descriptive Data for Resting, Reactivity and Recovery Conditions*

Measure	**Baseline**		**Maths**		**Singing**		**Recovery**	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
BPM *	66.20	9.30	78.05	11.29	79.10	14.18	71.02	10.23
Range	49 - 83		52 - 101		53 - 115		51 - 102	
IBI *	922.82	132.72	784.63	116.81	782.67	131.96	862.7	116.22
Range	723-1218		598-1135		517-1124		586-1169	
RMSSD*	57.24	37.06	43.97	31.00	40.43	23.94	66.08	48.29
Range	12.28 - 208.80		14.08 - 191.43		9.53 - 106.60		14.67 - 200.73	
HF *	929.50	1000.49	858.00	1525.25	606.33	669.72	1068.2	1306.20
Range	28 - 4048		37 - 9356		18 - 2733		22 - 5912	
Momentary stress/affect Range	1.45 1 - 5	0.80	2.97 1 - 5	1.23	3.16 1 - 5	1.18	1.95 1 - 5	1.08

Figure 4.2; *Mean BPM for Baseline, Maths, Singing and Recovery conditions.*

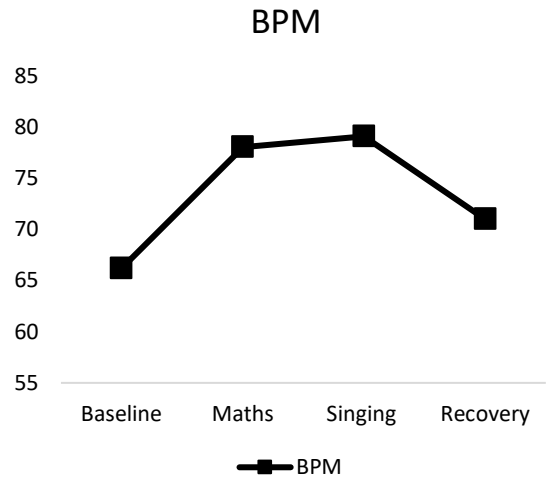


Figure 4.3: *Mean IBI for Baseline, Maths, Singing and Recovery conditions*

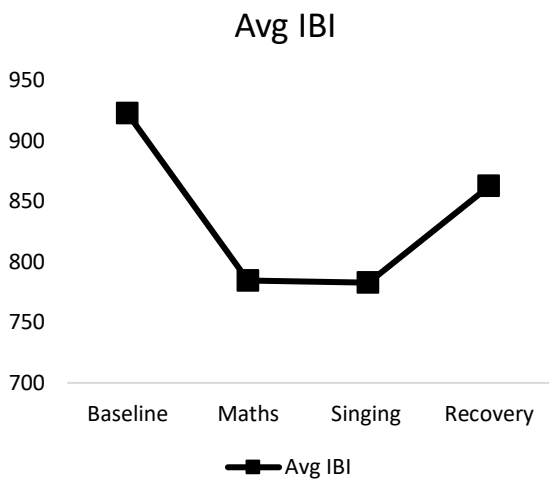


Figure 4.4: *Mean RMSSD for Baseline, Maths, Singing and Recovery conditions.*

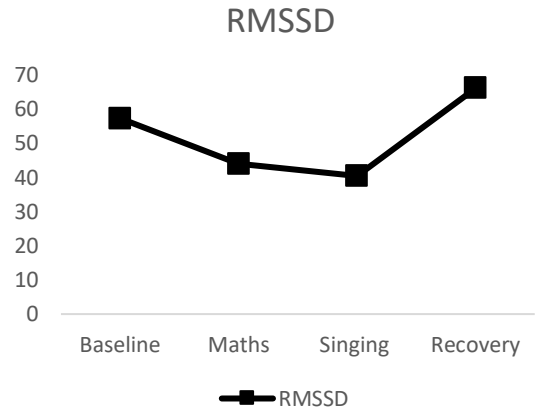
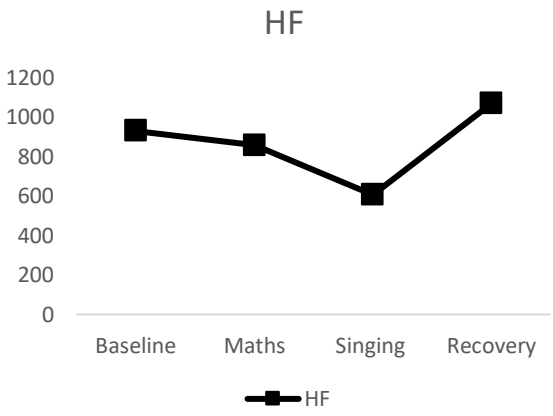


Figure 4.5 *Mean HF for Baseline, Maths, Singing and Recovery conditions.*



4.4.2. Beats Per Minute

A one-way repeated measures ANOVA was conducted to compare BPM during Baseline, Maths stress task, Sing-A-Song anticipation, and recovery. There was a significant overall effect for BPM, Wilks' Lambda = .28, F (3, 37) = 32.55, p < .000, multivariate partial eta squared = .725. Post hoc tests using a Bonferroni correction showed differences from baseline BPM (M= 66.2, SD=9.30) and (i) maths condition for BPM (Mean difference = 11.85, p < .000; (ii) Sing a song condition BPM (Mean difference = 12.90, p <.000) (iii) Recovery (Mean difference = 4.825, p < .000). Differences were also detected between recovery BPM (M=71.03, SD=10.24) and (i) maths condition for BPM (Mean difference = 7.03, p < .000; (ii) Sing a song condition BPM (Mean difference = 8.08, p < .000)

No statistically significant differences were found for BPM between the two stressor conditions.

4.4.3 Inter-Beat Interval

A one-way repeated measures ANOVA was conducted to compare IBI during Baseline, Maths stress task, Sing-A-Song anticipation, and recovery. There was a significant effect for IBI, Wilks' Lambda = .27, F (3,37) = 33.82, p < .000, multivariate partial eta squared = .733 Post hoc tests using a Bonferroni correction showed differences from baseline (M=922.82, SD =132.71) and (i) maths condition for IBI (Mean difference = -138.2, p < .000 (ii) Sing a song condition IBI (Mean difference = -139.95, p < .000) (iii) Recovery (Mean difference = -60.09, p < .000). Differences were also detected between recovery IBI (M=862.74, SD=116.22) and (i) maths condition for IBI (Mean difference = -78.11, p < .000; (ii) Sing a song condition IBI (Mean difference = -79.86, p < .000). No statistically significant differences were found for IBI between the two stressor conditions.

4.4.4. RMSSD

A one-way repeated measures ANOVA was conducted to compare RMSSD during Baseline, Maths stress task, Sing-A-Song anticipation, and recovery. There was a significant effect for RMSSD, Wilks' Lambda = .552, F (3, 37) = 9.99, p < .000, multivariate partial eta squared = .448 Post hoc tests using a Bonferroni correction showed differences from baseline RMSSD (M= 57.24, SD=37.06) and (i) Sing a song condition RMSSD (Mean difference = -16.79, p = .002). Differences were also detected between recovery RMSSD (M= 66.08, SD=48.28) and Sing a song condition RMSSD (Mean difference = -25.64, p = .001). No statistically significant differences were found for RMSSD between baseline and maths task or recovery and maths task.

4.4.5. High Frequency

A one-way repeated measures ANOVA was conducted to compare the HF component of HR during Baseline, Maths stress task, Sing-a-Song anticipation, and recovery. There was an overall significant effect for HF, Wilks' Lambda = .724, F (3, 37) = 4.698, p = .007, multivariate partial eta squared = .276 Post hoc tests using a Bonferroni correction showed significant differences between the sing-a-song condition (M= 606.33, SD=669.71) and (i) Baseline condition for BPM (Mean difference = - 323.18, p = .017; (ii) Recovery condition (Mean difference = 461.96, p < .036). No significant differences in HF were found between the Maths task and baseline or recovery.

4.4.6. Gender differences

Analysis of HR revealed mean differences for males and females. Independent-samples t-tests were conducted to compare mean scores for each component of HR for males and females across resting, maths, sing-a-song and recovery conditions. There were no significant

differences in BPM, IBI, RMSSD and HF scores across all conditions. It should be noted however that at baseline, on average males had lower mean BPM (M = 63.76, SD = 9.68) than females (M = 68.95 SD = 8.07) although this did not reach statistical significance (t (41) = 1.859, p = .071, two-tailed). At baseline males also had, on average, higher IBI (M = 959.86, SD = 146.58) and females (M = 881.4, SD = 101.75) although this did not reach statistical significance (t (41) = 1.981, p = .07, two-tailed).

4.4.7. Self-report-affect

A one-way repeated measures ANOVA was conducted to compare self-report stress during Baseline, Maths stress task, Sing-A-Song anticipation, and recovery. There was a significant overall effect of time, Wilks' Lambda = .38, F (3, 39) = 21.23, p < .000, multivariate partial eta squared = .62. Post hoc tests using a Bonferroni correction showed differences from baseline (M=1.45, SD =80) and (i) maths condition (Mean difference = 1.52, p < .000; (ii) Sing a song condition (Mean difference = 1.71, p < .000) and (iii) Recovery (Mean difference = .50, p = .039). No statistically significant differences were found for Self-report affect between the two stressor conditions. Post hoc tests using a Bonferroni correction also showed differences between baseline (M=1.45, SD=.80) and (i) recovery (Mean difference = .50, p = .039

4.5 Discussion

This study examined a stress protocol that used an anticipatory sing-a-song task as a task in combination with a maths subtraction task for inducing mental stress while adhering to a resting/reactivity/recovery design suitable for HRV analysis. Findings indicate that the protocol is effective in generating a stress response in laboratory. Statistically significant differences in HR measures and in self-report stress measures between baseline and stressor conditions were detected and the peak response was the anticipatory singing task.

Analysis of cardiovascular measures indicate significant differences between baseline and stressor conditions for BPM, IBI, RMSSD and HF. Significant within subject differences were detected between baseline and stressor conditions for BPM, with greater magnitude differences between baseline and the sing-a-song task. A main effect was also detected for IBI, with differences between baseline and both stressor conditions. Overall differences between baseline stressor conditions were detected in RMSSD, although post hoc analysis showed significant differences between baseline and singing task only. HF also showed a main effect for the combined stressor tasks but was not significantly different between baseline and maths task. The finding that RMSSD and HF only differed between baseline and singing task is interesting given the expectation that the maths task, which combines both evaluative stress and a greater cognitive load, would have a higher magnitude pattern than the anticipatory singing task. Given that RMSSD is reflective of vagally mediated changes (Thayer & Lane, 2000), that reflect parasympathetic activity and is correlated with HF, this could indicate that the anticipatory singing task, given its singular social evaluative task component, is reflective of a social interaction, while the cognitive elements may be contributing to increased BPM and decreased IBI. This underlines the importance of combining CV measures when examining elements of stressor tasks and also when considering what type of procedure to employ in

laboratory stress task. No significant differences were detected on any measure between the stressor conditions. In alignment with the physiological data, participants reported significant differences in self reports affective state for both maths task and anticipatory singing task. No differences were detected between the stressor conditions.

Brouwer and Hogervorst (2014) cite challenges including movement or body positioning (standing) and speaking as possible confounds in the TSST and have demonstrated the potential of the SSST in addressing these. The methodology in the current study, utilising both a maths task and anticipatory components that can be completed while in a seated, steady postural position, can reduce the likelihood of measures being affected by speaking and standing. In addition to evidencing the efficacy of the protocol, the current study has demonstrated that the procedure is robust when adapted to a resting/reactivity/recovery outlay. The adjustments to the procedure align the task with recommendations for experimental planning, data analysis and data reporting for HR type studies using cardiovascular measures to indicate psychobiological reactivity and recovery (Laborde et al., 2017; Quintana & Heathers, 2014). These include standardised baseline measurements with equivalence of time periods for stressor conditions and enable sufficient duration to allow calculation and comparison of both time domain and frequency components. Despite changes, the anticipatory singing task remained efficacious in combination with the maths task. This indicates its potential for use within future studies, that comprise a variety of tasks with various cognitive and evaluative elements. Allen et al, (2017) described requirement for replacement elements or standalone tasks that address habituation to repeated use of the TSST.

While the proof of concept has been demonstrated, there are limitations with the current study. It provides initial evidence of the efficacy of the singing task in combination with the math

task, but this has not directly been examined in the singing task alone. In addition, the tasks are not counterbalanced, this is related to the element of deception in the singing task, namely that singing will be required when it is not and the stressors could not be counterbalanced if required. This could be addressed in future study if required by actually asking participants to sing, which could constitute another stressor measurement period. An alternative strategy is to revise the procedure so that as part of the instructions given in the anticipatory phase, participants are told that only half of the participants were required to sing, and this will be decided at random, and at end of preparatory period. Doing this would not require any participant to perform the task, thus retain the time and resource advantages, protect the future efficacy of the task, as participants would not presume they will not actually be required to sing in future tasks, and reduce the likelihood of participants presuming other tasks will not be necessary, and so retains the researcher-participant trust as well as integrity of the stress challenge. It may also serve to enhance the stressor, as unpredictability has been reported to amplify stress (Peters, McEwen, & Friston, 2017). Future research using the anticipation only variant of the S-S-S-T should consider this task in combination with other stressor task, using different cognitive load and evaluative elements and in relation to individual difference factors which may amplify or ameliorate the magnitude of response to anticipated singing.

Despite these limitation, the findings remain robust as standard Trier procedures include dual stressor tasks routinely (Allen et al., 2017; Birkett, 2011; Brigitte M Kudielka et al., 2007b) and are sufficient for stress induction procedures. In addition, the dual stressor with combined time or five minutes allow matching of the ten-minute baseline and recovery periods. The conclusion of the current study is that the S-S-S-T carries distinct advantages, that include lower resource demands, reduced participant time burden and retention of the social evaluative, and cognitive performance challenges that are crucial components of effective stressors. In addition to these advantages, the S-S-S-T variant tested in this study is efficacious evidenced

by the non-invasive, sensitive measures of cardiovascular activity. This study was intended as a proof of concept and has demonstrated the efficacy of an anticipatory singing task as a lab stressor, and as a task replacement to the TSST with consideration for adherence to HRV measurement.

Chapter 5: Study three- Examining the use of an online adaptation of the Day Reconstruction Method.

5.1 Abstract

Research that uses physiological data collected in ambulatory contexts requires contextualised mapping of experience to match affective and psychosocial variables. Typically, sampling follows a number of methodological approaches; ecological momentary assessments (EMA) experience sampling methods (ESM) and retrospective diary type measures such as the day reconstruction method (DRM). Challenges exist for all collection methods such as adherence and the ability to integrate with other momentary and retrospective platforms. This study aims to demonstrate an adapted version of the Day Reconstruction Method that uses an online collection method, to examine adherence and compare the data to previous paper-based methods. Findings, indicate that episodic and affective data from this study align with findings reported for previous studies, with the inclusion of timestamped data that allows tracing of ambulatory adherence patterns. The collection method a robust means of documenting structured retrospective data and given the availably of timestamped data allows for the potential integration with momentary collection methods. The implications for data collection using this method and the advantages of including adherence data are detailed and further discussed.

Outputs

At the time of completion of this thesis this study is completed but has not been submitted for publication.

Contributions

The lead researcher and author on this study is Owen Jump. The work was supported and reviewed by Samantha Dockray

5.2 Introduction

In order to confidently relate physiological measures collected in ambulatory settings to emotional state, experience and specific contexts, researchers need to be able to be sure they have accurately and precisely recorded event time-frames and can link this to experiential data, including measures of affect and psychosocial experience and processes (Shiffman et al., 2008). Experiential data have been used to explore the relationships of psychosocial experience and context with a myriad of outcomes, including physical health and disease (e.g. Smyth, 2003), psychopathologies including anxiety, mood disorders and substance use disorders (Raugh, Chapman, Bartolomeo, Gonzalez, & Strauss, 2019) (Walz et al., 2014), and substance use disorders (Bertz et al., 2018). Experiential data must be sufficiently granular to allow the physiological data to have precision, such as exact and accurate time of collection, or by experience or context type, for example, where is the person, who are they with and what are they doing, but the gathering of this data should not place undue burden on the participants. Typically, sampling in real world contexts follow a number of methodological approaches; ecological momentary assessments (EMA) (Stone & Shiffman, 1994) experience sampling methods (ESM) (Csikszentmihalyi & Larson, 2014), and retrospective diary type measures. EMAs' are designed to focus on the participants current state and sample the person's experience (Shiffman, Stone, & Hufford, 2008) in their ecological context throughout the day, and involve repeated measurements. These can be synchronous or asynchronous, or anchored to specific events or experiences of the person. Retrospective diary type measures, such as the Day Reconstruction Method (DRM) (Kahneman, Krueger, Schkade, et al., 2004), involve prompts for the person to recall their experiences and emotions over a previous specified period, which may be a day, or 24 hour day/night period, and are completed at the end of the specific period of interest. EMA and the diary type measures have benefits, but the selection of methods must be guided by the research question and the requirements of the data, as well

as some more pragmatic concerns relating to researcher and participant burden, or the measurement concerns, such as recall bias.

Ecological Momentary Assessment (EMA) are methods designed to collect ecologically valid data, while participants go about their daily lives (Shiffman et al., 2008). The sampling strategy is designed to capture measures of participants current mood, psychological state, and behaviours for specified time-points or as general random sampling, which are selected dependant on the specifics of the research, and this may be synchronised to event, or timepoint. The resultant data represents a longitudinal, within subject capture of participant affect, psychological state, or any specified measure that can be self-reported, that can be then be used to describe the participant's affect, context, experience and behaviour, and this in turn, can be used to examine specific discrete periods, or used in a time-series analysis to map changes in person, over time. These EMA measures can also be mapped to biological measures, including to examine anticipatory and lagged effects of experience on biology, and vice versa. Data captured by EMA methods enable accurate time point associations of affect and behaviour to biological measures. The instrumentation may vary dependant on the specific research question, particularly in relation to what and type of psychological or affective measures are being assessed, for example measures related to affective (Kim et al., 2013) or health behaviours (Heron et al., 2017), and whether these are in research or clinical contexts (Trull & Ebner-Priemer, 2009).

Advances in technology, and in particular mobile and wearable devices have enabled a wider range of EMA platforms to be developed and allow more nuanced and targeted protocols that capture both biological and affective data to be implemented (Heron et al., 2017). Proponents of EMA platforms cite their ability to mitigate challenges associated with retrospective designs,

such as recall bias, and that situationally and temporally sited measurements provide more accurate representations of participant state, and so also enhancement of analyses that examine how these may change over time (Burke et al., 2017). However, certain types of EMAs' carry disadvantages, for example some measures can be complex and time consuming to complete, such as where multiple measurement periods in combination with reminders are sent to participants may do so at inconvenient times. This has implications regarding feasibility of use with particular populations, for example adolescents (Wenze & Miller, 2010) and can impact adherence (Moeller et al., 2014). Finally, the data yielded from multiple collection periods can have high resource for data processing and analysis (Smyth, 2003).

An alternative approach to capturing momentary experiential data is to use retrospective or recall measures such as interviews or diary type strategies that ask participants to recall the period of interest. Participants can be asked to report affect, situational data such as what they were doing and who they were with, in order to reconstruct their experience. Although retrospective measures have certain advantages over EMAs, such as no disruption of everyday activities and assessment of contiguous activates over the whole day (Kahneman, Krueger, Schkade, et al., 2004), they have been criticised for not addressing bias and error associated with recall (Wilhelm & Grossman, 2010). One measure designed to address the disadvantages of EMA and retrospective type measures, is the Day Reconstruction Method (DRM) (Kahneman, Krueger, Schkade, et al., 2004). The DRM is a hybrid measure that was designed to reduce recall bias by categorising experience into episodes and providing associated affective and psychological ratings to reconstruct participant experience more accurately. There are a number of advantages of the DRM, including, reduced disruption of normal daily activities, while offering the reflective experience of the diary practice that can produce rich

and detailed data, and gives participants the opportunity to report salient information (Kahneman, Krueger, Schkade, et al., 2004).

With regards to which measure is preferable, the type of sampling method employed should be informed by the research question, for example, whether sampling of data from specific time periods is required (such home/leisure contexts) or if whole day aggregate data is required. In cases where detailed differentiation of time periods is sought, timed stamped electronic methods are more desirable. Where whole day aggregate data are required or more general contexts the DRM may be considered. In each case the inclusion/exclusion criteria for each measurement period will have to be justified in relation to the research question and considered for suitability for participants. Approaches that allow integrative approaches of momentary and retrospective type measures are desirable for a number of reasons; for example, affect captured close to or at the same time an event occurs may be representative of a particular context and time. When participants have the opportunity to reflect on an episode, their retrospective appraisals can shape reporting of affect. For example, Kahneman demonstrated that pain experienced during a colonoscopy was differential reported depending on the immediate context (Redelmeier & Kahneman, 1996). This speaks to the question of whether the experienced self or the remembered self represents a truer reflection of experience (Kahneman & Riis, 2005). While two approaches have previously been viewed as dichotomous in research contexts, a number of studies have sought to examine the associations of the two types of measurement (Dockray et al., 2010; Kim et al., 2013), and they can also be conceptualised as complementary and necessary to accurately encapsulate a true picture of experience (Wilhelm & Grossman, 2010). This represents a more holistic approach to experience sampling and more accurately reflects therapeutic applications while maintaining the integrity of the data for research purposes. However, a number of methodological issues require consideration to allow

better integration of these measures. For example how do the limitations in accessing adherence to time points and response patterns in reconstructive studies place restrictions on how the these measures can be integrated with EMA measures (Heron et al., 2017; Kim et al., 2013; Wen et al., 2017). Diener (2014) suggests that study examining these factors as well as participant burden, time required to complete, and degree of adherence are desirable. While advocatess of the DRM cite its' ability to lessen burden and the lower the possibility of introducing experiential confounds, at least compared to other diary measures (Stone et al., 2006), time point adherence is difficult to determine (Diener & Tay, 2014) and the instrument can produce complex paper data records leading to high researcher burden. Although electronic based methods do exist and have increased in use, particularly in HCI contexts many researchers have constructed unique versions of the DRM in individual studies. The original DRM instrumentation is presented and completed in pen and paper format in the home or other ecological context, away from the researcher. Participants are responsible for recording the time, date and other pertinent and required details. Participants are guided by a sense of the value of their contribution, and strive for accuracy, as indicated by studies using saliva sampling time verification (e.g Kudielka et al 2012) but objective verification of the accuracy of the time or day of completion of the measure is restricted in terms of how participant activity and task completion can be observed by the researcher.

There is a requirement to adapt or develop a measure that will address the concerns of recall bias, precision of time, context and experience recording and burden, for both participant and researcher, and that enables this data to be smoothly integrated with other data on human experience and health, for example, cardiovascular activity, sleep and biomarkers associated with experience. integrative approaches. One approach to these challenges is to demonstrate the utility of an online version on the DRM. This study aims to address a number of these

challenges including, adaptation of the DRM to an online collection methodology, demonstrating that the use of the online version of the DRM matches the episodic and affective data reported for previous studies, and finally to demonstrate how the inclusion of timestamped data can be used to determine adherence patters for the DRM. The amendments to the methodology are documented and the returned data are detailed and further discussed.

5.3 Method

5.3.1. Participants

Participants (n=29, Male= 13, Female = 16, aged 19-65 Mean = 34.21 SD=12.12) were recruited from the general population using convenience sampling. Calls for participation were advertised using social media and university-wide email list. Participants who expressed interest were provided with an information sheet and were scheduled for a briefing session before a measurement session.

5.3.2. Design

This study uses a between-within subjects' design to evaluate an adaptation of Day Reconstruction Method approach. Original data were collected over a four-day measurement period using an electronic version of the DRM. This novel data was then compared with data from existing studies. Adherence patterns were examined using time-stamped data and are presented.

5.3.3. Materials and measures

Day Reconstruction Method- Electronic (DRM-E).

The Day Reconstruction Method (Kahneman, Krueger, Schkade, et al., 2004) is a measure used for assessing experiences and affect in daily life, and uses retrospective methods using a diary compromised of set time periods to assess participant experience and affect during the previous day. The DRM has been validated using other experience sampling assessments (Dockray, Grant, Stone, Kahneman, Wardle, & Steptoe, 2010) and has been shown to be a reliable method of monitoring affect and experience over defined epochs. The original DRM instrumentation consists of 4 'packets' of measures, that is completed by pen and paper. In the current study the packets were adapted to an online presentation, with sections presented contingent on

earlier answers, the self-report, demographic, and employment questions as well as the daily online elements were retained. The individual packets were adapted as follows,

Packet 1 of the DRM records general demographic information including age, gender education, marital status, living situation, ethnicity and income. Participants are asked to self-report on overall life satisfaction and general affect in contexts such as at home and at work. Participants completed this in pen and paper format while attending the briefing session, and no amendments to the original DRM were made.

Packet 2 develops the structured personal daily diary that participants are asked to construct during the ambulatory collection phase. Packet 2 provides a reflective space for the participants to reconstruct their episode list for the previous day. IN the original DRM structure, this packet is separated from other packets to maintain the integrity of the autobiographical recall, and was also completed using pen and paper, to encourage the self-reflection as in the original DRM format.

Packet 3 is completed online and asks participants to elaborate in further detail for each episode including, detail on the title of the episode, where they were, who they interacted with and self-report affect for each episode.

Packet 4 in the original DRM is a hybrid of affect scales for the data collection period and general employment information. For this study, daily questions are extracted, such as those related to daily mood and included them in the online daily diary measures. Affective states were captured on 7-point (0-6) Likert scales including, happiness, tiredness, worry and feelings of being hassled, angry and frustrated. The employment information was completed at the briefing attendance phase.

5.3.4 Procedure

Briefing session

Participants were contacted via email, after initial expression of interest and informed of the nature of the study and asked to attend a briefing session. Attendance was scheduled to enable the measurement of two work and two non-workdays for each participant. Briefing sessions were conducted in a research room and after consent procedures participants were asked to complete the DRM Packet one measures, including demographic and employment data. Participants were then provided with Packets two and three along with the daily questions from Packet four to be completed as part of the daily ecological setting procedure. In order to ensure accurate completion of the DRM in the ambulatory phase participants were given standardised instructions and completed a practice version of the DRM measure with the researcher during the briefing session. Participants were assigned a unique identification number for anonymisation purposes, and this number was used in all online phases to ensure anonymity. It should be noted that as in similar studies using the DRM (e.g. Dockray et al., 2010) this study was part of a larger study examining heart rate in ambulatory contexts.

Daily diary procedure

Participants were instructed to set aside time and sit in a quiet room alone and complete Packet two - the pen and paper self-reflective activity. Participants were informed that the diary section is completely confidential. Diary information for the day of completion represents entries for the previous 24-hour period. Packet two is presented as morning/afternoon/evening sections with time spaces for each. Each day period is further divided into episodes by the participant. These episodic data are then used to fill out the DRM Packet three daily measures. Packet three represents the online DRM task. Packet three comprises further detailed episodic data outlined in Packet two, including affective scales. The standard procedure for packet three

of the DRM requires this section to be completed in pen and paper format. This task was completed online for the current study. The data were entered by participants using Google forms. Each participant completed the four-day measurement period.

5.3.5. Data analysis

Data were collected using an online survey format and once data collection was complete, was uploaded to SSPS (Version 22). Data were assessed for errors in entry, such as entry of data on the wrong day, and categorised for adherence as, Fully Adherent, Partially Adherent and Non-adherent, prior to data analysis. Further detail of the adherence criteria are outlined in the findings. Demographic data and total number of episodes are reported. For episodic data the activity type, time of day of episode and type of social interaction are reported.

In addition, affective data are reported and grouped by activity for the four-day collection period. In order to assess if the affective data were consistent with previous studies the data were correlated with data from published studies using a Pearson product-moment correlation and the data for the analysis are also reported.

5.4 Findings

5.4.1. Demographic data

Participants (n=29, Male= 13, Female = 16, aged 19-65 years, Mean age = 34.21 SD=12.12) were employed as follows; Office Administrative (6), Service industry (5), Technical/research (11), Full time student (4), Health care worker (2), and Other (Archaeologist) (1). As a group, participants reported 926 total episodes (mean= 8.3 per day, sd = 4.4, Range 1-22) as follows; waking n=103, commute n=144, work n=109, socialising n=48, exercising, relaxing n=67, prayer/meditation n=11, phone/skype n=13, on the internet n=31, watching TV n=71, housework n=42, hobby n=10, family time n=20, intimate relations n=9, self-care n=29, eating n=96, driving n=4, shopping n=20, study n=9, bedtime routine n=17 and other n=17

Participants were instructed to collect data on two consecutive workdays and two consecutive non-workdays however when the data were analysed the following split for work/non work days 63% workdays 37 % non- workdays. A one-way between-groups analysis of variance was conducted to explore the impact of work/non workdays on both overall positive and negative affect for all episodes. Although there was a statistically significant difference in positive affect scores for work/non workdays, F (1, 925) = 25.50, $p < .05$. the actual difference in mean scores between the days was small and the effect size, calculated using eta squared, was >.02. There was also statistically significant difference in positive affect scores for work/non workdays, F (1, 925) = 12.40, $p < .05$. Similar to positive affect, despite reaching statistical significance, the actual difference in mean scores between the days was small and effect size, calculated using eta squared, was >.01. Given the small differences for affective the data reported for the following results are reported for all days and not parsed according to work/non workdays.

5.4.2 Episodic data

Table 5.1: *Total episodes reported by Activity, Social interaction and social interaction*

Total episodes reported by Activity type/Time of day/Social interaction						
Activity Type	N	Mean Hrs	**Time of Day**		**Type of social interaction**	
Waking	103	0.93	06:00-07:00	33	Alone	324
Commute	144	2.06*	07:00-08:00	59	Friends	85
Work	109	6.9*	08:00-09:00	75	Relatives	82
Socialising	48	2.36	09:00-10:00	64	Spouse/SO	193
Exercising	55	1.07	10:00-11:00	57	Children	27
Relaxing	67	2.17	11:00-12:00	50	Spouse & Children	20
Prayer/Med	11	1.14	12:00-13:00	61	Clients/ Customers	22
Phone/Skype	13	0.98	13:00-14:00	48	Co-Workers	79
Internet	31	1.70	14:00-15:00	51	Boss	7
Watching TV	71	2.28	15:00-16:00	49	Students	12
Housework	42	1.46	16:00-17:00	34	Pets	4
Hobby	10	2.88	17:00-18:00	54	Other	70
Family time	20	2.27	18:00-19:00	55		
Intimate Relations	9	0.52	19:00-20:00	71		
Self-care	29	0.86	20:00-21:00	47		
Eating	96	0.90	21:00-22:00	44		
Driving	4	1.31	22:00-23:00	27		
Shopping	20	1.42	23:00-00:00	22		
Study	9	2.87	00:00-06:00	25		
Bed	17	6.00				
Other	17	2.01				
Total	925					

**Note that mean "work" and "commute" times are calculated per work-day measurement period. All other means are representative of full 4-day period.*

Table 5.1 shows the number of episodes, grouped by activity, time of day and social interaction for the four-day collection period. The results show that the most frequently reported activity was commuting to and from work/college. This was followed by working, waking/morning routine, and eating. Table 5.1 also details the time of day for reported episodes. The largest number of episodes were reported during the 8.00-9.00 time period (N=75), followed by 19.00-20.00 (N=71) and 9.00-10.00 (N=64). Data are also reported for the type of social interaction participants were engaged in during each episode. The participants reported being alone the majority of episodes (N=324), followed by interactions with their Spouse/SO (N=193), and Friends (N=85).

5.4.3. Affective Data

Table 5.2 shows the mean affect, grouped by activity for the four-day collection period. The results show that overall positive affect (Happy, Competent, Warm/Friendly and Enjoying Myself) have higher mean scores (Mean= 3.25, range = 2.56 – 4.00) than negative affect (Impatient, frustrated/annoyed, depressed/blue, hassled/pushed around, angry/hostile, worried/anxious and criticised/put down) (Mean 0.59, range = 0.30 – 1.12) Tired is not included in aggregate negative affect scores. The highest positive mean affect scores were reported for Intimate relations (N=9), Socialising (N=48), and Family time (N=20). The highest Negative affect scores were returned for Work (N=109), Commute (N=144), and Study (N=9).

A one-way between-groups analysis of variance was conducted to explore the impact of work/non workdays on both overall positive and negative affect for all episodes There was a statistically significant difference in positive affect scores for work/non workdays, F (1, 925) = 25.50, $p < .05$. However, despite reaching statistical significance, the actual difference in mean scores between the days was small and the effect size, calculated using eta squared, was >.02. There was also statistically significant difference in positive affect scores for work/non workdays, F (1, 925) = 12.40, $p < .05$. However, similar to positive affect, despite reaching statistical significance, the actual difference in mean scores between the days was small and effect size, calculated using eta squared, was >.01

The relationship between episode type and positive affect scores for the sample in the current study was compared to published data from the original DRM study (Kahneman, Krueger, Schkade, et al., 2004) using a Pearson product-moment correlation coefficient. This analysis was limited to the original episodes reported by Kahneman, et al (2004) including: Commute, Work, Socialising, Exercise, Relaxing, Prayer/Meditation, Phone/Skype, Watching TV,

Housework, Family time, Intimate Relations, Eating, Internet use, and Shopping. There was a strong positive correlation between positive affect, $r = –.79$, $n = 14$, $p < .05$, reported in both studies. The relationship between episode type and negative affect was also investigated using a Pearson product-moment correlation coefficient. There was a strong positive correlation between negative affect, $r = –.65$, $n = 14$, $p = .01$, reported in both studies.

Table 5.2: *Mean affect rating by episode type*

Mean affect rating per episode type														
Affect description	Happy	Competent	Warm/ friendly	Enjoying Myself	Overall Positive	Impatient	Frustrated/ Annoyed	Depressed/ Blue	Hassled/ Pushed	Angry/ Hostile	Worried/ Anxious	Criticised/ Put down	Tired	Overall Negative
Mean score	3.32	3.38	3.28	3.04	**3.25**	1.30	0.86	0.36	0.27	0.24	1.00	0.10	2.13	**0.59**
Activity														
Waking	3.09	3.41	3.11	2.73	**3.09**	1.26	1.04	0.54	0.50	0.46	0.99	0.13	2.96	**0.70**
Commute	2.84	3.26	2.80	2.31	**2.80**	2.32	1.25	0.56	0.44	0.43	1.36	0.21	2.23	**0.94**
Work	2.66	3.49	3.03	2.18	**2.84**	2.89	1.55	0.48	0.63	0.43	1.14	0.11	1.91	**1.03**
Socialising	3.95	3.54	3.95	4.06	**3.88**	0.60	0.43	0.06	0.33	0.25	0.60	0.10	1.68	**0.34**
Exercising	3.68	3.86	3.43	3.56	**3.63**	1.07	0.41	0.29	0.11	0.07	0.84	0.07	1.64	**0.41**
Relaxing	3.59	3.31	3.21	3.48	**3.40**	0.42	0.42	0.29	0.10	0.09	0.85	0.03	2.64	**0.31**
Prayer/Med	3.16	3.33	3.08	3.08	**3.16**	1.50	0.08	0.66	0.58	0.08	0.58	0.25	1.90	**0.53**
Phone/Skype	3.15	2.69	3.46	3.00	**3.08**	1.07	1.07	0.46	0.15	0.30	1.15	0.23	2.15	**0.63**
Internet	2.87	3.29	2.77	2.45	**2.85**	1.48	1.38	0.54	0.25	0.03	1.29	0.00	2.35	**0.71**
Watching TV	3.47	3.32	3.16	3.60	**3.39**	0.53	0.78	0.50	0.04	0.47	1.23	0.14	2.00	**0.53**
Housework	3.29	3.58	3.21	2.78	**3.22**	2.04	0.97	0.46	0.21	0.21	0.58	0.17	1.39	**0.66**
Hobby	3.40	3.50	3.62	3.81	**3.58**	0.72	0.81	0.14	0.00	0.22	0.83	0.00	0.61	**0.39**
Family time	4.00	3.80	3.80	3.65	**3.81**	0.4	1.25	0.20	0.05	0.05	0.75	0.00	1.7	**0.39**
Intimate Rel	4.11	3.77	4.11	4.00	**4.00**	0.55	0.55	0.66	0.44	0.22	0.88	0.33	1.22	**0.52**
Self-care	3.16	3.23	2.80	2.56	**2.94**	1.43	0.86	0.53	0.26	0.23	1.46	0.03	2.13	**0.69**
Eating	3.62	3.50	3.43	3.36	**3.48**	0.46	0.44	0.27	0.18	0.16	0.78	0.09	1.49	**0.34**
Driving	2.75	2.75	2.75	2.00	**2.56**	1.75	1.00	0.00	0.00	0.00	1.75	0.00	2.75	**0.64**
Shopping	3.70	3.65	3.6	3.4	**3.59**	1.35	0.65	0.10	0.30	0.15	0.65	0.00	1.75	**0.46**
Study	3.27	3.45	3.63	2.63	**3.25**	2.18	0.81	0.45	0.27	0.09	1.81	0.00	3.27	**0.80**
Bedtime	3.31	3.00	3.15	3.26	**3.18**	0.63	0.42	0.31	0.00	0.31	0.31	0.15	3.15	**0.30**
Other	2.71	3.33	2.71	2.04	**2.70**	2.66	2.09	0.14	0.76	0.85	1.19	0.14	1.61	**1.12**

Table 5.3: *Mean affect rating by social interaction*

Mean Affective rating per social interaction													
Affect description	Happy	Competent	Warm/ friendly	Enjoying Myself	Overall Positive	Impatient	Frustrated/ Annoyed	Depressed/ Blue	Hassled/ Pushed Around	Angry/ Hostile	Worried/ Anxious	Criticised/ Put down	Tired	Overall Negative
Mean score	3.22	3.56	3.35	3.05	**3.30**	1.53	1.01	0.33	0.43	0.30	0.93	0.11	2.04	**0.66**
Type of social interaction														
Alone	3.17	3.24	2.69	2.44	**2.89**	1.57	0.90	0.41	0.28	0.28	1.03	0.12	2.37	**0.66**
Friends	3.54	3.48	3.77	3.70	**3.62**	1.14	0.48	0.18	0.15	0.17	0.62	0.07	1.94	**0.40**
Relatives	3.07	3.39	3.17	3.13	**3.19**	1.07	1.17	0.54	0.26	0.59	1.09	0.14	2.23	**0.69**
Spouse/SO	3.55	3.61	3.76	3.60	**3.63**	0.76	0.78	0.45	0.26	0.19	1.20	0.10	1.88	**0.53**
Children	3.48	4.03	3.66	3.11	**3.57**	1.59	0.88	0.33	0.29	0.18	0.55	0.00	2.40	**0.55**
So&Child	3.10	2.90	4.05	3.40	**3.36**	0.75	1.10	0.25	0.25	0.45	0.35	0.05	0.55	**0.46**
Clients	3.27	3.72	3.40	2.50	**3.22**	3.00	1.54	0.18	0.36	0.40	0.68	0.04	2.40	**0.89**
Co-Workers	2.82	3.34	2.96	2.37	**2.87**	2.41	1.37	0.67	0.51	0.32	1.05	0.08	1.70	**0.92**
Boss	2.71	3.42	2.85	2.00	**2.75**	2.14	2.14	0.71	2.00	0.42	2.14	0.57	2.74	**1.45**
Students	3.08	3.50	3.66	2.75	**3.25**	2.00	0.58	0.16	0.41	0.08	1.33	0.00	2.50	**0.65**
Pets	3.75	4.50	3.25	4.50	**4**	0.25	0.25	0	0.	0	0	0	2.25	**0.07**
Other	3.10	3.60	3.00	3.12	**3.21**	1.62	0.98	0.19	0.40	0.53	1.10	0.19	1.53	**0.72**

Table 5.3 shows the mean affect by for each type of social interaction for the four-day collection period. The results show that overall positive affect (Happy, Competent, warm/friendly and enjoying myself) have higher mean scores (M=3.20, Range 2.89- 4) than negative affect (Impatient, frustrated/annoyed, depressed/blue, hassled/pushed around, angry/hostile, worried/anxious and criticised/put down) (M=0.64, range= 0.40-1.45) Tired is not included with the aggregate negative affect scores. The highest positive affect scores were reported while participants interacted with their pets (N=4), Spouse (N= 193), Friends (N=85), and their Children (n=47). Highest negative affect scores were reported for interactions with their Boss (N=7), Co-Workers (N=79), and Clients/Customers (N=22).

The relationship between mean scores social interaction and positive affect for current study and the original DRM study (Kahneman, Krueger, Schkade, et al., 2004) study was investigated using a Pearson product-moment correlation coefficient. This analysis was limited to the original episodes reported by Kahneman, et al (2004) including: Alone, Friends, Relatives, Spouse/SO, Children, Clients/Customers, Co-Workers, and Boss. There was a strong positive correlation between positive affect for social interaction, $r = –.84$, $n = 8$, $p < .01$, reported in both studies. The relationship between episode type and negative affect was also investigated using a Pearson product-moment correlation coefficient. There was a strong positive correlation between negative affect and social interaction, $r = –.93$, $n = 8$, $p = .01$, reported in both studies.

Table 5.4: *Mean affect rating by time of day*

Mean Affective rating per Time of Day

Affect description	Happy	Competent	Warm/ friendly	Enjoying Myself	**Overall Positive**	Impatient	Frustrated/ Annoyed	Depressed/ Blue	Hassled/ Pushed Around	Angry/ Hostile	Worried/ Anxious	Criticised/ Put down	Tired	**Overall Negative**
Mean score	3.25	3.39	3.17	2.95	**3.20**	1.45	0.94	0.42	0.27	0.28	1.00	0.11	2.21	**0.64**
Time of Day														
06:00-07:00	2.92	3.08	2.68	2.64	**2.83**	2.00	1.12	0.4	0.16	0.16	1.04	0.08	3.48	**0.71**
07:00-08:00	3.06	3.47	3.11	2.42	**3.02**	1.62	1.01	0.52	0.30	0.15	1.18	0.08	2.50	**0.69**
08:00-09:00	3.04	3.44	3.06	2.68	**3.06**	1.85	1.13	0.38	0.61	0.42	1.18	0.29	2.27	**0.84**
09:00-10:00	3.06	3.30	3.30	2.75	**3.10**	1.63	1.13	0.45	0.41	0.31	0.91	0.02	1.64	**0.69**
10:00-11:00	3.23	3.63	3.07	3.00	**3.23**	1.61	0.96	0.32	0.40	0.25	1.00	0.09	1.86	**0.66**
11:00-12:00	3.40	3.34	3.42	3.10	**3.32**	1.24	0.84	0.40	0.48	0.30	1.22	0.12	1.58	**0.66**
12:00-13:00	3.43	3.43	3.33	3.16	**3.34**	1.20	0.82	0.43	0.28	0.26	1.23	0.11	1.66	**0.62**
13:00-14:00	3.19	3.56	3.17	2.85	**3.19**	1.58	0.90	0.27	0.27	0.29	0.92	0.04	1.52	**0.61**
14:00-15:00	3.24	3.59	3.31	2.88	**3.26**	1.86	1.35	0.53	0.47	0.57	1.16	0.20	1.65	**0.88**
15:00-16:00	3.31	3.39	3.16	3.00	**3.22**	1.63	0.86	0.24	0.31	0.22	0.94	0.04	1.98	**0.61**
16:00-17:00	3.26	3.44	3.26	2.71	**3.17**	1.97	1.24	0.62	0.21	0.26	1.09	0.06	2.12	**0.78**
17:00-18:00	3.37	3.48	3.22	2.96	**3.26**	1.26	0.89	0.30	0.35	0.39	0.80	0.17	1.48	**0.59**
18:00-19:00	3.49	3.58	3.13	3.09	**3.32**	1.02	0.58	0.47	0.18	0.24	0.89	0.18	1.95	**0.51**
19:00-20:00	3.44	3.39	3.24	3.39	**3.37**	1.27	0.90	0.32	0.20	0.41	0.96	0.13	2.20	**0.60**
20:00-21:00	3.49	3.30	3.19	3.26	**3.31**	0.85	0.45	0.43	0.11	0.17	0.77	0.15	2.30	**0.42**
21:00-22:00	3.43	3.43	3.16	3.30	**3.33**	0.77	0.84	0.41	0.16	0.18	1.09	0.00	2.57	**0.49**
22:00-23:00	3.41	3.19	3.44	3.37	**3.35**	0.78	0.70	0.41	0.07	0.07	0.78	0.04	2.63	**0.41**
23:00-00:00	3.15	3.25	3.25	2.90	**3.14**	1.55	1.10	0.75	0.00	0.60	0.85	0.25	3.20	**0.73**
00:00-06:00	2.92	3.08	2.68	2.64	**2.83**	2.00	1.12	0.40	0.16	0.16	1.04	0.08	3.48	**0.71**

Table 5.4 shows the mean affect, grouped by time of day for the four-day collection period. The results show that overall positive affect (Happy, Competent, warm/friendly and enjoying myself) (Mean = 3.20, range 2.83-3.37) have higher scores than negative affect (Impatient, frustrated/annoyed, depressed/blue, hassled/pushed around, angry/hostile, worried/anxious and criticised/put down) (Mean = 0.64, range 0.41- .88). Tired is not included in negative affect scores.

Affect- time of day

Figure 5.1: *Time of day by positive affect rating*

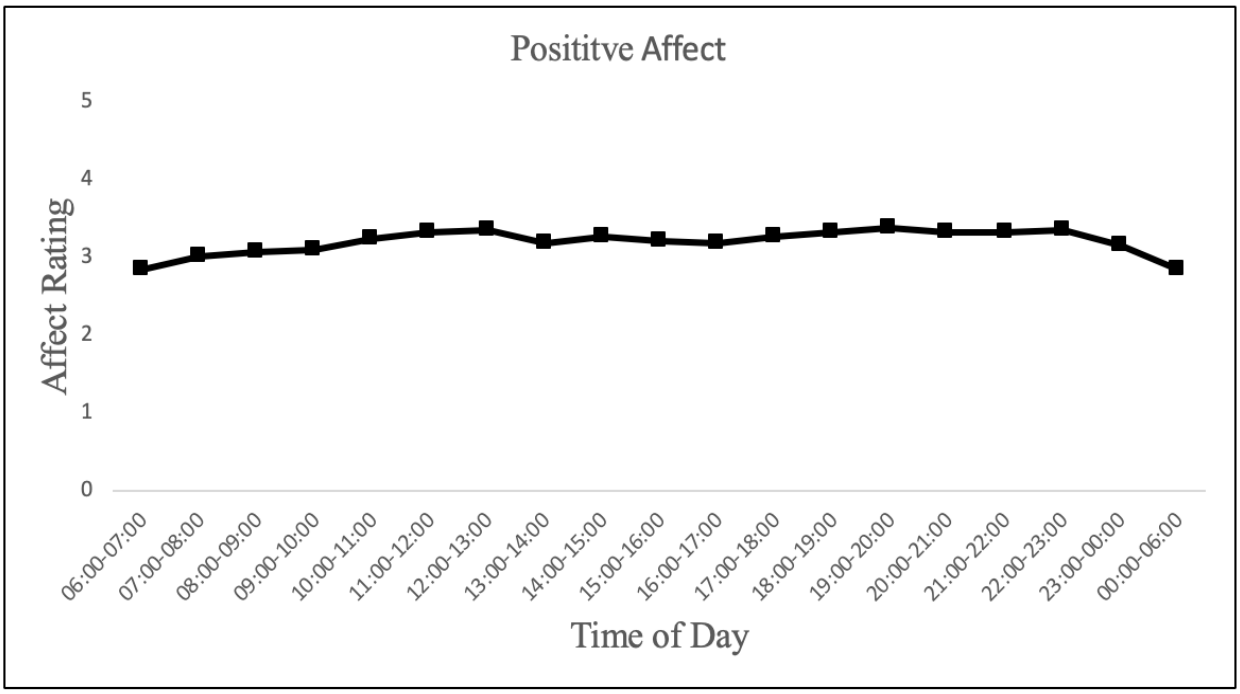


Figure 5.1 shows positive affect for time-of-day data aggregated for the four-day collection period. Aggregated affect ratings for all days including both work/non-workdays are shown. Positive affect includes the combined data for Happy, Competent, War/Friendly, and Enjoying myself. The data shows the diurnal pattern of positive affect, with lowest positive affect scores reported in the morning and late night.

Figure 5.2: *Time of day by negative affect rating*

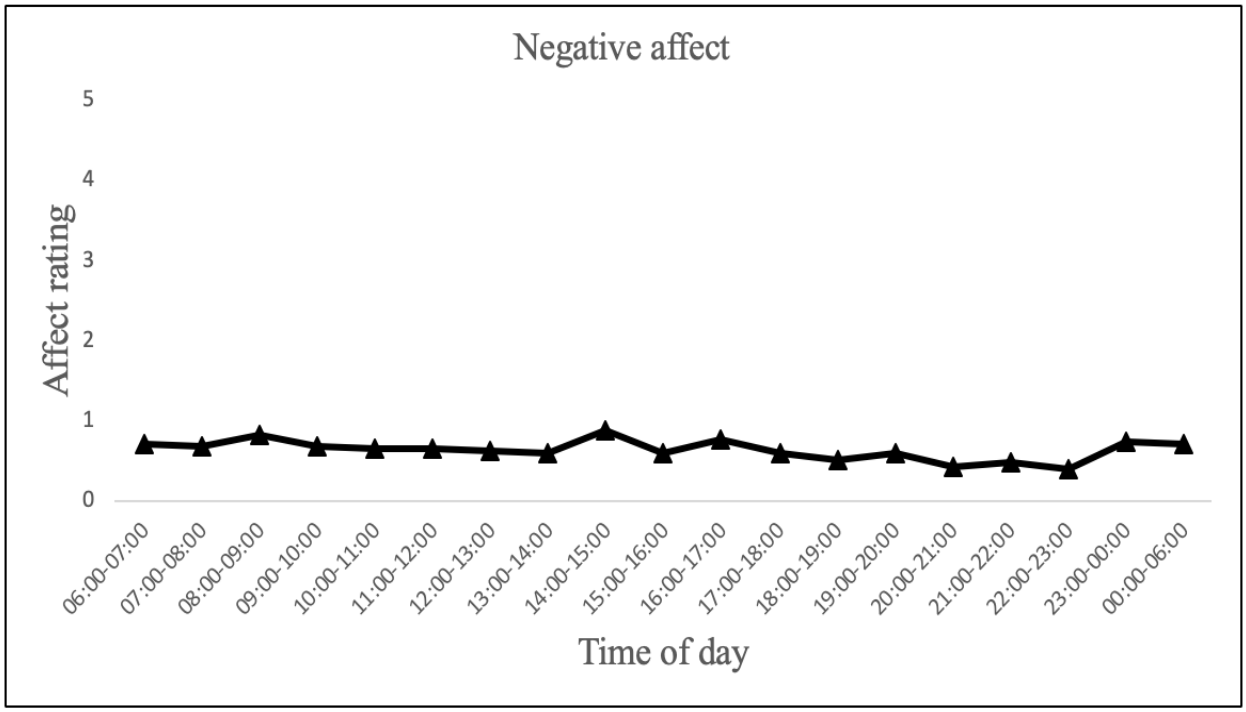


Figure 5.2 shows negative affect for time-of-day data aggregated for the four-day collection period. Affect ratings for both work/non-workdays are shown. Positive affect includes the combined data for Happy, Competent, War/Friendly, and Enjoying myself. The data shows the diurnal pattern of positive affect, with lowest positive affect scores reported in the morning and late night.

Figure 5.3: *Time of day by tiredness rating*

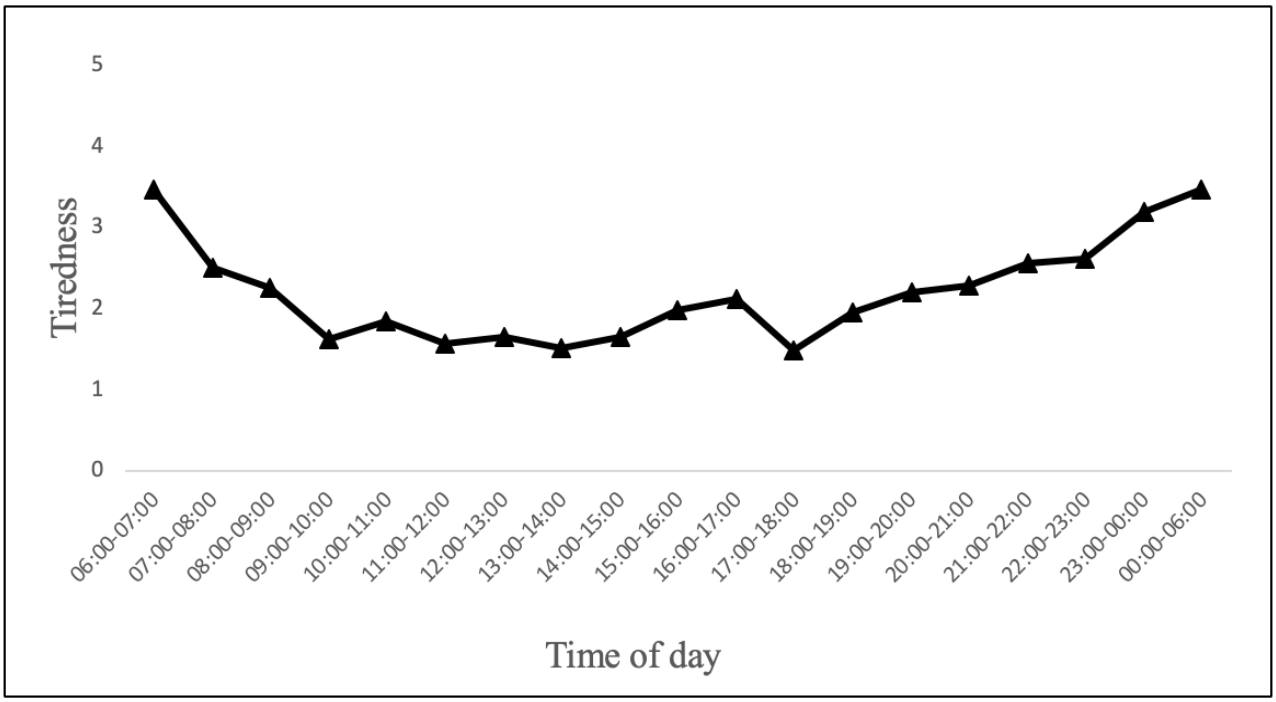


Figure 5.3 shows self- tiredness for time-of-day data aggregated for the four-day collection period. Affect ratings for both work/non-workdays are shown. The data shows the diurnal pattern of tiredness with the highest scores reported early in the morning and late at night. It is also noted that higher levels of tiredness are reported in mid-afternoon, with lowest reported tiredness at mid-morning and early evening.

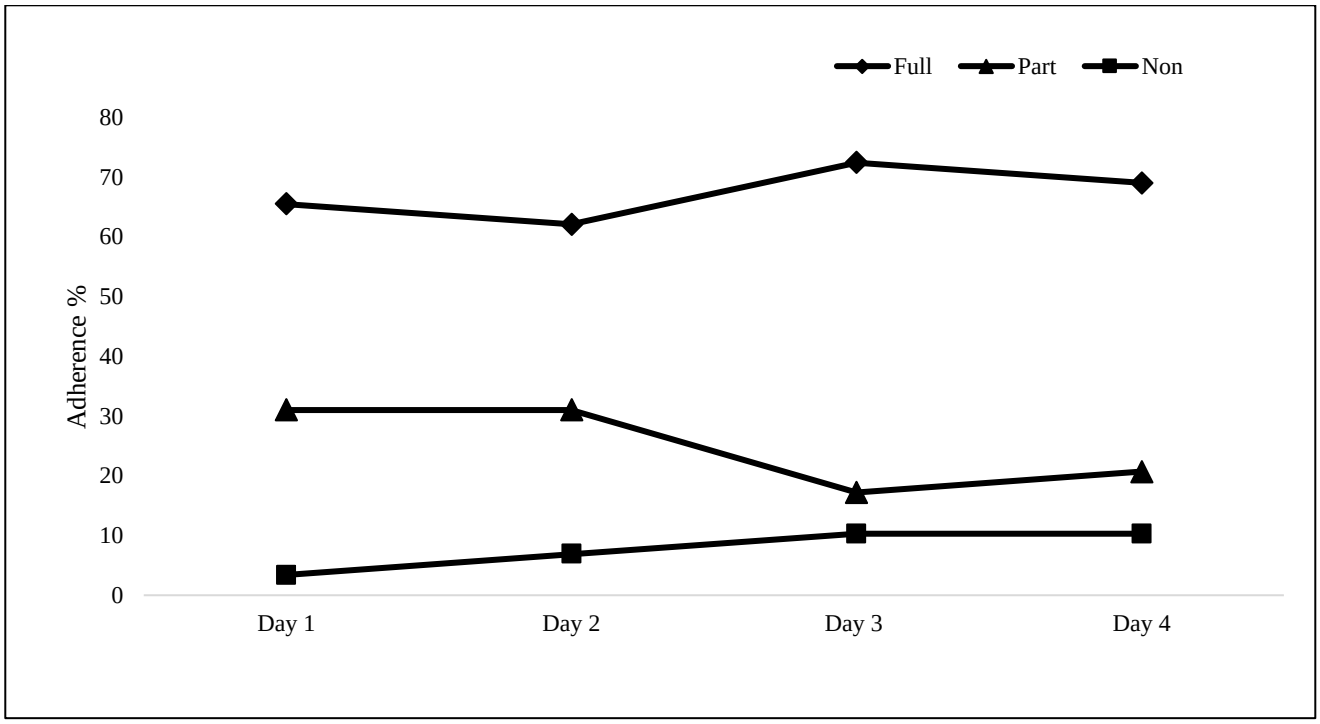
5.4.4 Adherence

Overall adherence

Timestamped data were collected as participants accessed the website and entered data. Adherence was determined by assessing if the data entry was completed on the appropriate date and within the corresponding evening timeframe, as directed by the DRM instructions. Adherence data was categorised as follows; Fully Adherent as all episodes were fully completed within the designated timeframe, Partially Adherent, when the designated day was correct, but not the designated time frame (e.g. evening) and Incomplete, is data were not entered for a day. Eleven (38 %) participants maintained full adherence throughout data collection, 14 (48%) of participants had Partial Adherence, and these deviations from instructions were primarily where all episode data from a day were entered at once. Several (N=4; 13%) of participants were categorised as non-complete.

Daily adherence

Figure 5.4: *Daily adherence percentages, categorised as Full, Part and Non-completion*



Daily adherence was also examined to determine if patterns of completion were influenced by the day of entry. Figure 5.4 shows the daily adherence patterns by day of instructed completion. Day one data indicated 65.5% full adherence, 31% partial/late entry and 3.5% non-entry. Day two data indicated 62.1% full adherence, 31% partial/late entry and 6.9% non-entry. Day three data showed 72.4% full adherence, 17.2% partial/late entry and 10.3% non-entry Day four data showed 69% full adherence, 20.7% partial/late entry and 10.3% non-entry.

5.5 Discussion

5.5.1. Overview

This study presented here is an adaptation of the DRM with changes applied to the original methodology including, migration to an online collection method, amendments to the structure of the DRM instrumentation to suit online data completion, and maintenance of the integrity of episodic and affective data while implementing changes to the protocol. The study addressed a number of challenges associated with the existing DRM, including confirmation of adherence to timing of measurement, and a reduction in methodological resource load that occurs with the original format of the DRM method. Overall, the results provide evidence that the online completion of the DRM has greater ability to determine adherence, provides the opportunity for researchers to engage with participants during data collection, and potentially increases total complete data. The findings and study details, including reporting patters for the DRM, adherence, episodic data, and affective patterns are further discussed.

5.5.2. Adherence

Adherence patterns, and in particular the ability to time stamp data represents a major advantage of using electronic collection methods as presented here. To our knowledge, adherence to daily protocols has not been demonstrated with the DRM. Timestamped data were captured and compared to the timepoints participants were instructed to complete data entry at laboratory attendance. While other study such as Dockray, et al (2010) and Daly et al (2010) describe the use of electronic surveys to collect DRM data, adherence patterns are not elaborated. A categorization for adherence is suggested namely, Full, Partial and Non completion. The majority of participants in this study were full of partially adherent with a small number deemed non-adherent. The ability to time stamp data carries a number of distinct advantages, ones to our knowledge have not been described in relation to the DRM previously.

These include the ability to ensure data entry is carried out at the reported time, and this study demonstrated that non-adherence to measures is relatively low, with the majority of our participants being fully or partially adherent. Following that, the inability to interrogate collection patterns is improved using the presented protocol because the level of granularity provided by the timestamped data. This adds a level of analysis that can be used to screen data for partially or non-adherent participants. The ability to discriminate between partial and fully adherent participants is advantageous because it allows for the inclusion or exclusion, not only of certain participants from analysis, but also of certain collection days for individual participants, where they previously would have needed to be excluded entirely.

Another advantage is real-time online verification of adherence, where previously there had been no oversight possible once participants begin an ambulatory DRAM collection phase. This advantage may directly contribute to reducing partial non-adherence, allowing researchers to initiate reminder strategies. The protocol presented here allows for real time interaction with the data, and therefore contact with the participant to query adherence and offer additional guidance is possible. Additionally, participants can be informed at the instructional phase that their data are time stamped and this could influence adherence. Describing verification checks has been shown to enhance adherence, for measures such as saliva sampling as well as records of mood and experience (Dockray, O'Neill & Jump, 2019).

The final advantage with the presented methodology is that with a better understanding of adherence using timestamped data, retrospective diary type measures can be more reliably integrated with other experience sampling measures. This is desirable for a number of reasons, which are firstly methodological, where data from momentary and retrospective patterns can be interrogated together by researchers providing better insight into episodic and affective

patters. The question of whether capturing the experiencing self or the remembered self (Zajchowski et al., 2017) is the best method of interrogating experience remains, however this may be a false dichotomy, and that experience sampling, if it to be truly reflective of real-world psychological phenomena and affect will need to integrate the two. As sampling accuracy and the integration of methods improves this question can be further explored.

5.5.3. Similarities of data between online DRM method and paper-based methods

Overall, the comparison of data collected using online means with that collected in paper forms are suggestive of similarity in data returned, at least compared by episode number, and affect reporting. Participants in this study reported data for the full range of episodes captured by the DRM. Commuting represented the most frequently reported episode, followed by working, waking, and eating, with more episode reported in the morning time followed by early evening. The lowest period of episode reporting was during mid-afternoon, and in particular workdays, these data are similar to those previously reported for both DRM and ESM, at the 15:00 to 16:00 timepoint (Dockray, et al, 2010). Although Kahneman suggests in the original DRM study (Kahneman, Krueger, Schkade, et al., 2004) that reporting patterns of work or leisure episodes are influenced by time pressure, this is not further elaborated upon. Other authors have suggested reporting patterns may be a characteristic of the salience, with higher reporting being related to higher variability in experience (Stone et al., 2005). The data for this study returned similar mean scores to those reported Kahneman et al (2004) in the original DRM study, and also a sample using random sampling of weighted episodes (Anusic et al., 2017), although the number of episodes were lower than those reported by (Schwarz, & Stone, 2004) and slightly higher than other studies (Dockray et al., 2010; Lee et al., 2017). Our analysis indicates that the original DRM and our sample aligned very well for type of episode and social interaction.

The current study details affective patterns as they relate to episode type and also social interaction, diurnal patterns of affect across the day, and tiredness. Our study examined affective data in relation those reported by Kahneman (2004), and Anusic (2017), and examined affect by type of episode and social interaction. The findings indicate that these data correlate with those from existing studies, and this is an indicator that the adaptation of the collection method does not seem to influence the overall pattern of affect or episode reporting suggesting that it is impossible to say with certainty that the mode of collection has any effect of the numbers of episodes or affect ratings without conducting a repeated time-periods study.

This study also describes patterns of positive and negative affect, and again our data are reflective of other studies that use the DRM, where reporting of positive affect is more prevalent than negative affect (Möwisch, Schmiedek, Richter, & Brose, 2019; Anusic, Lucas, & Donnellan, 2017; Mellor-Marsá et al., 2016; Dockray et al., 2010; Schwarz, & Stone, 2004). It should be noted that patterns of affect did not align with Kahnemans' study in relation to diurnal affective patterns, where that study reports more variation across the day, although it should be noted they only report on negative affect and tiredness. Our daily patterns follow more closely those described by Anusic in a larger scale study (Anusic et al., 2017), where positive affect reporting is dominant and negative affect is reported to a much lesser degree. Another similarity to our study was that overall lower positive affect scores were reported to be lower in comparison to those in Kahnemans study. These patterns are also reflected in other studies using the DRM in European samples (Mellor-Marsá et al., 2016; Möwisch et al., 2019). Our data, although taken from a much smaller sample, matches the affective scoring and diurnal patterns of the European samples more closely. The differences observed may be due to cultural differences in reporting or inherent in the study design themselves and further study is required to explore this.

Relationships between social interaction indicated strong correlations for both positive and negative affect, suggesting that the type of social interaction was highly salient for the participants. The patterns reported here for social interaction were mostly in agreement with previous study but again do show some differences. Kahneman (Kahneman, Krueger, Schkade, et al., 2004) reported the highest affect scores for interaction with their friends, followed by relatives, with their S/O third. Although largely in agreement with our data, Kahneman reports lower affect ratings for interaction with ones' children, with this interaction only slightly above interactions with a boss or co-workers. Anusic (Anusic et al., 2017) reports patterns of affect in social interaction that indicate friends, followed by relatives and then children, returned the highest positive affect scores also return similar affect scores. Although this study adapted the methodology of the DRM the data related to social interactions and subsequent affective patterns are in agreement with previous study.

As mentioned previously, the current study also details diurnal rhythms of positive and negative affect, and tiredness. In our sample positive affect remained relatively stable throughout the day, with limited variation in overall scores. Similar patterns were observed in negative affect, although tiredness followed a characteristic diurnal pattern with the highest reported scores in the early morning and evening, with a small peak during late afternoon. Again, the diurnal patterns of affect observed in our study also more closely align with the data reported in the European context. Although only negative affect and tiredness are reported by Kahneman, our data indicate much flatter affective patterns throughout the day, which aligns more closely with the other European studies, with similar agreement regarding patterns of tiredness.

5.5.4. Episode weighting and data analysis

In a review of the DRM Diener (Diener & Tay, 2014) highlights the need for further work to understand the degree to which time weighting of episodes influences the returned scores for affect. Both the original DRM (Kahneman, Krueger, Schkade, et al., 2004) and a following study (Stone et al., 2006) do not report on time weighted episodes. Some later study suggests that weighting is necessary (Dockray et al., 2010), and better correlations between affect for EMA and the DRM are present if data were attenuated for time spent on task. Others address similar questions but do not employ weighting to their analysis (Kim et al., 2013) and report similar data to the original DRM and the Stone study. Our findings indicate the relationships remain robust despite the non-weighting of episodes in the current study. The decision to not employ weighted means in the current study was taken due to the small sample size and predominance of positive affect. A recent study (Lee et al., 2017) has proposed even more complex analysis should be applied, with aggregated affect, that should also account for activity type and social interaction. Given these considerations, and the potential need for precision in reporting adherence, and potentially accurate time stamping, online methods offer significant advantage to paper based methods, and it is encouraging that the relationships between affect and episode/social situation remain.

5.5.5. Limitations and Future study

This study provides initial evidence for how online DRM collection methods might support adherence. Although this study does not specifically detail our ability to 'spot-check' adherence, or that timestamps could be examined closely, it is possible that participants realised the researchers could actually do so, and this may have led to greater adherence. This possible effect was not able to be assessed with our design, but other researchers may be able to test the effect of explicit statements about the capacity to monitor adherence and intervene to increase

engagement and reduce attrition. A further consideration is the nature of the DRM completion in the current study. Given that this study retained some of the pen and paper elements from the original DRM, in order to best model the reflective experience, there could be a time separation between the time of pen and paper completion and electronic data entry. A limitation of this study that should be addressed in further iterations is that these two activities are anchored together.

It is also noted that the data returned included a greater number of measures of workdays than non-workdays. The initial design set out to have equivalent numbers of work and non-workdays, especially important given previous reports of levels of affect on work and leisure days e.g. (Ryan et al 2010, Biskup, Kaplan, Bradley-Geist & Membere, 2019). This design was explicit in recruitment, and in the briefing given to participants, where participants confirmed their work pattern for the nominated days of measurement, yet preliminary analyses indicated that many designated non-workdays contained episodes of work. This may indicate the fluidity of work patterns across all days, or that in the current sample, people were in occupations that were more likely to have unanticipated work demands or offers of paid work, for example in retail or hospitality roles, or in occupations where is possibly more likely that some work is done on non-workdays (e.g., knowledge worker) (Ciolfi & Lockley, 2018), or the blurring of boundaries between work and leisure, what has been described as 'weisure' (Bartlett, 2014). The mix of work and non-work activities within a day must be recorded accurately to do any examination of affect, context and experience, and the episodic structure of the DRM approach enables this.

This study did not set out to directly address the questions of episode weighting or other analytical approaches that have been described in relational to momentary assessment, or the

DRM in particular, as our focus was on testing the online mode. There is a standing question in relation to experience, episode duration and affect which requires dedicated study and it is hoped that a future focused examination will be undertaken. Despite these limitations the methodology presented here does support the advantages of the collection method and data that may be collected by future studies, for example it allows the participant adherence to be interrogated. The ability to timestamp data and track adherence permits novel questions to be asked of the returned data, for example, do affective patterns interact with adherence? Does the length of the collection period matter in relation to adherence, i.e., does adherence change over time? How does tiredness and time taken to complete the measures effect adherence? While this study does not address these questions, the adapted methodology allows them to be examined in future study.

5.5.6. Conclusions

That idea that EMA and DRM type methodologies are dichotomous and in competition is anachronistic. EMA platforms that integrate retrospective measures in a structured way, and provide opportunity for reflective experience, such as the adapted, online DRM, will provide measures that are better able to account for human experience. This study has shown the potential to address a number of the challenges to integrating the DRM with EMA approaches and highlight the direction of future possible study aimed at addressing these challenges.

Chapter 6: Study Four - From the lab to the living room: Measuring heart rate variability in ecologically valid contexts

6.1Abstract

Cardiovascular activity is a frequently used measure in psychophysiological science, in both laboratory and controlled conditions, as well as ecological contexts. Models of human health that draw upon cardiovascular measures obtained in a controlled condition have an implicit presumption that these measures are good indicators of how a person may respond in ecological contexts. In order to study patterns of heart rate in laboratory contexts, researchers have measured tonic/resting, and phasic/reactive HRV in relation to standardised stress challenges and measurement protocols. However, advances in technology enables prolonged periods of measurement in ecological contexts, using discrete ECG capable devices. This allows full spectrum measurement of continuous cardiac output and monitoring of movement via accelerometer data that permit the use of HR measurement in ambulatory research designs. Although the precision of these data is still disputed, the transition to ambulatory measurement is accelerating and research that links the two is timely. This study details patterns measurement of ambulatory contexts, including experience sampling, and extends the measurement of participants' HRV across four full days of 24 hour-a-day measurement, and describes their concordance with laboratory contexts. Patterns of baseline and stressor HRV inside the laboratory are matched to ambulatory measurement epochs and group and fixed effect patterns of HRV change and are described across conditions. The findings indicate robust relationships between laboratory and ambulatory HRV and affect for BPM an RMSSD and a novel means of screening ambulatory data is described using affective patterns and experience sampling to associate it with laboratory measures for HRV type studies.

Outputs

At the time of completion of this thesis the study presented in chapter six is under review following a revise and resubmit as

Jump, O., & Dockray, S. (2021) From the lab to the living room: Measuring heart rate variability in ecologically valid contexts. Journal of psychophysiology (Under review)

Contributions

The lead researcher and author on this study is Owen Jump. The work was supported and reviewed by Samantha Dockray.

6.2 Introduction

Cardiovascular activity has been widely incorporated as a measure of psychophysiological reactivity and response in laboratory contexts (Laborde, Mosley, & Thayer, 2017; Quintana & Heathers, 2014), and it has been associated with a variety of health outcomes (Kemp & Quintana, 2013; Thayer, Åhs, Fredrikson, Sollers, & Wager, 2012), and psychosocial variables (Huang et al., 2013), with many studies using measures of heart rate reactivity, blood pressure, and heart rate variability (HRV). HRV is the set of measures used to assess change in heart rate across time and frequency domains over defined epochs (Quintana & Heathers, 2014). Time domain measures assess the variability in timing of the successive beats and is indicated by distance in R –R peak intervals, with variations in Inter-Beat Interval (IBI), Root Mean Square of Successive Differences (RMSSD), Standard Deviation of Non Normal intervals (SDNN) used as descriptions of variability (Fred Shaffer & Ginsberg, 2017). Frequency domain measures have been associated with physiological parameters for example, Ultra Low Frequency (ULF) has been linked to circadian rhythm and metabolic demands (Fred Shaffer et al., 2014), Very Low Frequency (VLF) has been linked to physical movement (Bernardi et al., 1996) and parasympathetic activation (Taylor et al., 1998). The Low Frequency (LF) band has been associated with baroreceptor activity and blood pressure (Fred Shaffer & Ginsberg, 2017) and proposed as an index of sympathetic outflow (Reyes del Paso et al., 2013), however some debate exists regarding the inclusion of LF as a reliable index of sympathetic activation (Billman, 2013). High Frequency (HF) components alone have been reliably associated with vagal tone and reflect parasympathetic activity (Porges, 2007). LF/HF ratio has also been proposed as an index of autonomic balance (ChuDuc et al., 2013; Quintana & Heathers, 2014). However as the usefulness and meaning of LF components continues to be scrutinised its' inclusion requires careful consideration and justification (Billman, 2013; Reyes del Paso et al.,

2013) and this also extends to the use of the LF\HF ratio as a marker of sympathetic/parasympathetic balance.

Both time domain and frequency based measures are associated with a wide range of psychosocial and affective correlates (Kreibig, 2010; W. Porges, 2001) and a number of theoretical frameworks have been proposed that seek to explain the functional relationships of HRV and psychobiological wellbeing. One is the Polyvagal theory (Porges, 1994, 2007), which focuses on the neural regulation of the heart, and outlines the mechanisms by which emotional arousal is indexed by HF components. This framework is used to describe how HRV indexes vagal tone and parasympathetic function (ChuDuc, NguyenPhan, & NguyenViet, 2013), and related social functioning (Porges, 2001;Quintana, Guastella, Outhred, Hickie, & Kemp, 2012). In addition, the Neurovisceral Integration model (NVI) (Thayer & Lane, 2000) describes how top down processes, mediated via the prefrontal cortex are linked to cardiovascular function and form the basis of individual differences in response to the environment, that are linked to psychosocial outcomes and cognitive function, such as attention and affect (Thayer & Fischer, 2013). Both the Polyvagal Theory and the NVI model allow for specificity to behavioural and psychological variables in bio-behavioural research and have been widely used in laboratory contexts.

In order to study patterns of heart rate in laboratory contexts, researchers have often used standardized stress tasks and measurement protocols and measured tonic/resting, and phasic/reactive HRV in relation to challenge. Tasks and activities used to study HRV in laboratory contexts include cognitive tasks, public speaking tasks, stress interview, emotion induction and social evaluative stress experiences (Chida & Hamer, 2008). Tasks such as the Trier Social Stress Test (Kirschbaum, Pirke, & Hellhammer, 1993; Kudielka, Hellhammer, &

Kirschbaum, 2007) have become the gold standard in laboratory paradigms, using combined cognitive and social evaluation based stressors and afford high replication value in experimental stress induction (Allen et al., 2017). These tasks have been employed to elucidate different responses in cardiovascular activation (Allen, Kennedy, Cryan, Dinan, & Clarke, 2014), including HRV patterns, and have been linked to pathophysiological processes brought about by sustained hypertensions and allostatic load (Thayer, Yamamoto, & Brosschot, 2010), and have also been associated with psychopathologies such as, depression (Burke, Davis, Otte, & Mohr, 2005; Young, Lopez, Murphy-Weinberg, Watson, & Akil, 2000) and anxiety (Gerra et al., 2000; Young, Abelson, & Cameron, 2004).

Although standardised laboratory stress tasks are widely used, criticisms of the strategy exist, including that they have narrow task representation (Wetherell & Carter, 2014), habituation to stressors (Schommer, Hellhammer, & Kirschbaum, 2003; Arvidson, Sjörs, & Jonsdottir, 2017), and ecological validity (Allen et al., 2017) and subsequently researchers have sought methodologies that more closely model real-world stressors. These types of studies (Thomas. Kamarack & William R. Lovallo, 2003) have included specified tasks with differential stimuli that can be employed depending on the research question (Wetherell et al., 2017) and physiological measures such as CVR and HRV, either in aggregate across longer time periods, or at intervals throughout the day, or in relation to specified tasks at discrete time periods. Jump & Dockray (Jump & Dockray, 2020) describe a stress protocol using an anticipatory sing-a-song task in combination with a maths subtraction task to address ecological validity. The protocol has advantages as it addresses confounds such as movement and speaking and is offered as a more ecologically valid laboratory protocol. The task has specificity to particular social contexts also and is therefore suitable for study in combination with ambulatory data.

Task specificity in the laboratory is an important consideration (Wetherell & Carter, 2014) as more reliable and ecologically valid techniques are required for observational research and clinical purposes, as standardised laboratory-based stressors may not represent a typical or frequently occurring stressor for the individual (Schwerdtfeger, Schienle, Leutgeb, & Rathner, 2014), where psychophysiological responses may also be mediated behaviourally, for example by avoidance of certain stressors, or changes in behaviours, for example, exercise or cigarette smoking. Studies that examine HR in both laboratory and ambulatory contexts, and their concordance, are few, and are not consistent in their findings. Dikecligil and Mujica-Parodi (2010) have described a comparison of laboratory stressors and an acute ambulatory stressor of completing a skydive and show that laboratory responses provide a marker for real world acute stressor reactivity, with good correlations between short term and longer term HRV. Hare and colleagues (Hare, Wetherell & Smith., 2013) examined habituation to a similar task using cortisol measures and report divergences between self-reported affect and biological measures in novice and experienced sky divers. Gerin et al (1993) describe how the reproducibility of blood pressure measurement in ambulatory contexts is moderately high if activity measures are standardised (Gerin et al., 1993). Kamarck et al, (2003) describe considerations for HR measurement and content validity across contexts in a later review (Kamarack & Lovallo, 2003). There are some inconsistencies in reports of a relationship between laboratory and ecological responses though, Schwerdfeger et al, (2014) studied CVR and affective states in laboratory and ambulatory contexts, specifically if strong reactivity (CVR) and laboratory baselines are related to ambulatory HR. Their findings indicate that ambulatory HR and laboratory CVR were not related, suggesting that laboratory reactivity might not be a reliable indicator of everyday CV activity (Schwerdtfeger, Schienle, Leutgeb, & Rathner, 2014). However, the lack of concordance was mediated by baseline measures, specifically where elevated baseline

measures and reactivity were detected, elevated ambulatory HR patterns were also present. In a further study (Schwerdtfeger & Dick, 2019) did find associations in attenuated HRV and resilience in ambulatory contexts. This suggests that rather than magnitude of response, individual patterns of HR that follow elevated baseline patterns are more susceptible to ambulatory stressors.

Further study is required to link laboratory and ambulatory patterns of reactivity however the challenge is to accurately define and measure experience in ecologically valid contexts, in ways that can be confidently related to responses to stressors in laboratory settings. Descriptions of HR profiles in laboratory, while informative, may only describe reactivity to acute or atypical, stressors, and while reactivity has been linked to pathophysiological processes, they do not fully describe the experience of daily stress that may be more relevant in models of individual differences in stress and psychobiological wellbeing. That is, individual vulnerability to stress can be profiled by measuring reactivity to acute type stressors but it is questionable whether they are experientially accurate to real world contexts.

Developments in technology have enabled prolonged, ambulatory measurement in ecological settings, for example, by using discrete ECG capable devices, such as Actiheart (Kristiansen et al., 2011; Takken et al., 2010). This allows full spectrum measurement of continuous cardiac output, and monitoring of movement via accelerometer data, that permit the use of HR measurement in ambulatory research designs (Laborde et al., 2017; Quintana & Heathers, 2014). In order to extend sampling outside of the laboratory researchers require accurate methods of capturing both physiological and experiential data, that can be reliably associated. Typically sampling of experience and affect in real world contexts follows two main methodological approaches; Ecological Momentary Assessments (EMA) (Stone &

Shiffman, 1994) that focus on the participants current state and sample in context throughout the day; and retrospective diary type measures such as the Day Reconstruction Method (DRM) (Kahneman, Krueger, Schkade, et al., 2004), that sample at the end of the day. While studies such as that by Schwerdfeger use EMA type data, retrospective type data such as that captured by the DRM can also be employed in these contexts. The DRM uses a structured daily reconstruction of participant experience in conjunction with specified episodes and affective scales to capture participant experience. The main advantage of the DRM is that it does not disrupt normal daily activities, and offers the reflective experience of the diary practice that can produce rich and detailed data, and it gives participants the opportunity to report salient information that can be missed by EMA measures (Kahneman, Krueger, Schkade, et al., 2004).

As the use of technologies aimed at the general public to record their own behaviour, experience, and physical functioning, for example, using of wearable devices and lifelogging apps, doubts have been expressed about the accuracy of these applications (Alexander, Minhajuddin & Joshi, 2017; Plante et al., 2016). Study linking these measures to verified laboratory paradigms is preferable both in terms of experiential measurers and physiology. Drury et al have proposed an integrative model of HRV and health, suggesting a "promising digital epidemiology, which can facilitate objective population health studies as well as clinical applications" (Drury et al., 2019). Ecologically valid HRV measures will contribute to this study of heart rate, however ambulatory monitoring challenges such as controlling confounds, such as movement, and the accurate definition and capture of experience require careful consideration. (Fred Shaffer & Ginsberg, 2017).

The current study examines patterns of cardiovascular activity in both laboratory and ambulatory contexts, using HRV measures and retrospective diary-type measures. The study

measures participants patterns of HRV in a controlled laboratory setting, employing a laboratory protocol outlined by Jump and Dockray (2020). This laboratory protocol is used as it is advantageous in terms of methodological resource load and has task specificity to allow comparison with ambulatory data. These laboratory data are examined in combination with data from a four full 24hr ambulatory period for each participant, to represent a range of contexts and behaviours. Participant affect is also captured in the laboratory and ambulatory contexts using the Day Reconstruction Method. These data are combined to sample ambulatory context across four days for each participant. Baseline and reactivity patterns in the laboratory will be examined in relation to ambulatory HR, specifically that HRV patterns in the laboratory can be matched to ambulatory contexts, given adequate definition and description of ambulatory conditions. Initially, laboratory baseline HRV patterns are assessed and compared to laboratory reactivity. Laboratory baseline is also compared to ambulatory resting/positive affect and negative affect conditions, and finally ambulatory conditions were examined in relation to laboratory reactivity. Our hypothesis extends to include, firstly, laboratory baseline and laboratory reactivity conditions will differ significantly, secondly, closely matched ambulatory patterns will not differ significantly from laboratory baseline, and finally, laboratory reactivity and ambulatory resting condition will differ significantly. The following presents inter (group) differences and intra (individual) patterns of HR for the variety of conditions.

6.3 Method

6.3.1. Participants

Participants (n=25 Male= 11, Female = 14, aged 19- 56, Mean 34, SD=10.90) were recruited from the general population using convenience sampling. Participation was advertised using social media and university email. Participants were contacted after initial expression of interest and given a link with surveys to complete and a schedule for research lab attendance.

6.3.2. Design

This study uses a within-subjects design to compare a laboratory-based stress task, including a standardized maths task followed by an anticipatory 'sing-a-song' task in resting / reactivity segments, and patterns HRV in matched ambulatory contexts.

6.3.3. Materials and Measures

Demographics

Demographic data were captured including, age, gender, employment, income, marital status and ethnicity. The required demographic data was informed by the Day Reconstruction Method (Kahneman, Krueger, Schwarz, et al., 2004). Information in relation to employment directly informed how each participant recorded and reconstructed their daily experience, and which days were scheduled for recording, therefore it was important to provide adequate record of these data to account not only for between group differences but also possible confounds.

Physiological measures

Physical characteristics of the participants were recorded in the laboratory, including, mass (KG), height (CM), age, and gender (Antelmi et al., 2004). Measures of height, mass were taken in the laboratory by the researcher. Participants reported physiological and lifestyle

factors (both stable and transient) that could cause potential confounds in the heart rate data in accordance with recommendations set out by Laborde et al , 2017 (Laborde et al., 2017), including; age (Umetani et al., 1998), gender (Koenig & Thayer, 2016) and alcohol use (Quintana et al., 2013). Participant height and mass were measured in the lab and BMI (Antelmi et al., 2004) was calculated. Participants self-reported medical conditions or medications prior to laboratory attendance, including; Cardio-active medications, antidepressants or anti-hypertensives (Cohen, 2001; Kemp et al., 2010; Schroeder et al., 2003). Participants self-reported estimated resting and maximum active heart rate. All of the above parameters are required by the Actiheart software provide data for confounds and group differences in HR profiles, such as those correlated with BMI.

Electrocardiogram measures

ECG measurement was carried out using the Actiheart measurement device (Brage, Brage, Franks, Ekelund, & Wareham, 2005). Actiheart is a validated (Kristiansen et al., 2011) non-invasive method of measuring physical activity and energy expenditure (PAEE). (Crouter et al., 2008). Actiheart is attached over the skin and can be used for long recording periods (up to 21 days) in an ambulatory setting. Actiheart is used to measure ECG heart rate and movement as follows,

Heart rate- Actiheart is ECG capable with a sensitivity of 250uV with a range of 30 bpm – 250bpm. ECG is detected and amplified and sampled at 128Hz. The device is capable of capturing and calculating full r-r wave intervals. Actiheart uses the resultant data in its' proprietary software to calculate and resolve HRV measures that are presented to the user. ECG data for defined epochs is calculated and a proprietary algorithm sorts and cleans data prior to presenting it to the user. R-R intervals for at least 16 heart beats are captured for each epoch and averaged. Values outside of a +/- 25% 8range are removed and the remaining data re-

averaged (CamNtech, 2018). This represents the data returned to the user. Additional quality indexes (range 0-1) for each epoch are returned and can be examined by the user and excluded if necessary.

Physical activity - The Actiheart device also returns measures of physical activity for the participant. The device measures movement in 'counts of unit per time' at an 8-bit resolution in a frequency range of 1hz-7hza and sampling rate of 32Hz. The device contains a piezo electric element that generates a measurable transient charge when subjected to acceleration. The resultant charge produces a voltage signal that is divided into a predefined range of positive and negative levels (- 128 to + 128). This range is sampled 32 times a second. This time/second measurement is then applied by the proprietary software to the user defined epoch (CamNtech, 2010, 2018) returning the activity measure for the epoch.

The combined measures of the Actiheart represent the participants movement and heart rate during a specified period. Information is downloaded via accompanying proprietary computer software to allow examination of heart rate.

Laboratory Stress task

The laboratory task employed in this study is an adaption of the sing-a song stress task (Brouwer & Hogervorst, 2014) adapted to adhere to a resting/reactivity/recovery outlay, suitable for assessing resting and tonic HRV(Jump & Dockray, 2020). The task assesses reactivity using a serial subtraction maths task and an anticipatory stressor and is described below.

Resting - Baseline

Baseline measures for the Actiheart that conform to standards outlined by Fishel (2007) (Fishel, Muth, & Hoover, 2007) were conducted for all participants. Participants remained seated and resting in a chair for a duration of 10 minutes. Participants are required not to move to reduce artefacts in their baseline heart rate data. Participants were not asked to keep their eyes closed but were asked not to move if possible. No instructions were given with regards to hand position. Measures of HR and movement were recorded via the Actiheart device during the baseline period.

Reactivity – Maths task

The maths task component is a computer-based maths task written and presented on PsycoPy Software (Peirce, 2019). Similar to the maths element of the standard Trier Social Stress Test (Kirschbaum et al., 1993), participants are required to complete a serial subtraction task while under social evaluation, provided by constant observation of a researcher. The maths task is presented to participants and they are told that they must achieve a fictitious score and the number of iterations of the test that they must achieve. This element is included to increase perceived external social evaluative threat. Scripted comments, such as "Most people have done much better by now" were made to the participant at predetermined times during the task to increase the evaluative element

Reactivity- Singing task

The singing phase of the reactivity section is an adaptation (Jump & Dockray, 2020) of the stressor procedure outlined by Brouwer & Hogervorst (2014) employing 'anticipation of singing' as a primary stressor as a variant of the 'talking' task used in the original Trier Social Stress Test (Kirschbaum et al., 1993). This includes an anticipatory stress phase induced by a

surprise introduction of song lyrics and an instruction to sing the song in public. The singing phase requires the participant to sit on a chair in front of a microphone and camera. The researcher is present, sitting at right angles, in close proximity to the participant as they are given time to prepare to sing. A minute-by-minute countdown is included – increasing to every fifteen seconds for the final minute- followed by a ten second countdown at the end of the anticipatory period. At the end of the anticipatory phase the participant is told that they are not required to sing.

Recovery phase

A five-minute recovery phase was completed by all participants. The recovery phase, which mirrored the baseline measurement was included to allow a secondary comparison to the task situation (Quintana & Heathers, 2014). Recovery consisted of five minutes of rest sitting in a chair in the laboratory while continuing to wear the ECG device. Participants are again instructed not to move to reduce artefacts in the HR data.

Self-report affect

A self-report affect scale was included at five periods of measurement during the laboratory task, including, pre-baseline, baseline, maths task, singing task, and recovery. The affect scale requires participants to self-report five dimensions of subjective effect on a rating of 1-5, including, measures of relaxation, stress, control and anxiety the self-report measures were included to allow for quantification of participant self-perceived affect during the task. These data were utilised later to extract the matched ambulatory phases for comparison.

Day reconstruction method – Electronic (DRM-E).

The Day Reconstruction Method (Kahneman, Krueger, Schkade, et al., 2004) is used for assessing experiences and affect in daily life. The technique uses retrospective reconstruction using a diary compromised of set time periods to assess participant activity and affect during the previous day. This measure has been included as it represents a less burdensome method of sampling daily experience and allows uninterrupted sampling without the introduction of experiential aspects which may act to confound ecological validity it. Although the DRM offers lower possibility of experiential interference, time point accuracy and daily adherence to diary completion can be compromised in DRM type studies (Diener & Tay, 2014). The DRM has been validated using experience sampling assessments and has been shown to be a reliable method of for monitoring affect and experience over defined epochs (Dockray, Grant, Stone, Kahneman, Wardle, & Steptoe, 2010). In the studies presented here the DRM was adapted and the daily elements were hosted online to allow for ease of data entry and monitoring of participation by the researcher.

Affective data were captured via the DRM scales for the ambulatory phase. Episodic data in the DRM provides scaled affective data for each episode. These data were used to inform the data analysis, and subsequently determined how physiological data were extracted and analysed.

Ambulatory heart rate measurement.

Ambulatory heart rate was captured via the Actiheart device. Actiheart allows for up to 21 days of uninterrupted 24-hour heart rate measurement and movement via accelerometer data. All participants completed four full days of 24-hour heart rate monitoring while wearing the Actiheart device. The data retrieved from the Actiheart device were subjected to analysis, informed by affective and episodic data returned by the participants

6.3.4. Procedure

Laboratory attendance

Participants attended the laboratory and were asked to refrain from caffeinated drinks and smoking for at least one hour prior to attendance. Participants were greeted in a friendly but neutral manner by a researcher dressed in a white laboratory coat. Participants were then escorted to the laboratory. Once the consent procedures were completed, the participants were asked to complete the self-report measures. Once consent procedures were completed the participants completed the laboratory tasks described above, including, baseline, maths stress task and anticipatory singing task. All participants successfully completed the stressor task procedure and a matched recovery period. The participates were then de-briefed regarding the laboratory procedure to ensure no prolonged reactivity in accordance with ethical requirements.

Ambulatory procedure

Once laboratory procedures were completed the participants were briefed on the requirements of the four-day ambulatory measurement period. Participants were refitted with the ECG device and instructed that the measurement period would begin the following day at 6am. This was to account for any confounds caused by the collection day and give 4 days of data collection. Participants were given instructions on how to remove the Actiheart device if required, and to change the patches adhering the device if required.

Participants then completed the four-day measurement period. No special instructions were given, and participants were asked to behave in a usual manor and not limit any activity due to monitoring.

The participants were also given instructions on how to complete the diary measures.

Diary measures were completed by the participants as required at the end of each collection day. Participants reconstructed the episodes for each collection phase and entered their data on

an online forum. Once each collection period was completed the participants returned the ECG device to the research team and the diary measures and cardiovascular data was subjected to analysis.

6.3.5. Data Analysis

HRV Data

HR data for this study were captured for four 24-hour periods for each participant via the ECG device and exported using the Actiheart proprietary software. The software returns measures of beats per minute (BPM), Inter-beat interval IBI (Maximum, minimum and average measures of IBI are returned), root mean of successive square differences (RMSSD), standard deviation of NN intervals (SDNN). Also, frequency domain measures of High Frequency (HF ms^2), Very Low Frequency (VLF), Low frequency (LF ms^2) and finally Low frequency/High frequency ratio (LF/HF ms^2) are captured. For the purposes of the current study, measures of BPM, IBI, RMSSD and HF components only are reported, this is owing to disagreement and reliability regarding the LF components. Actiheart returns ECG data for defined epochs and its' proprietary algorithm sorts and cleans data prior to presenting it to the user. Examination of movement is permitted using accelerometer data returned by the ECG device. HRV data for both the laboratory phases and ambulatory measurement periods were exported from the Actiheart proprietary software to the SPSS statistical package for analysis.

Affective data

Laboratory affective data were captured at five timepoints, including all laboratory conditions and each participant reported affect for their multiple episodes for each of the four-days form their ambulatory period. A full report of these data is presented in the results section. Affective data in the laboratory were recorded using momentary affect scales. Affective data in

ambulatory contexts was recorded using episodic affect scales in the DRM. These data were grouped and mean scores for positive and negative affect were calculated. The ambulatory data were assessed and used to contextualise and match the laboratory and ambulatory physiological data as follows. Firstly, laboratory affect was examined to assess responses across conditions in the laboratory. Secondly, pre-determined cut off points, where affective data, for both negative and positive conditions with an upper and lower threshold, were used decide inclusion for ambulatory affect.

Ambulatory conditions criteria

In order to match participant data to laboratory conditions a number of pre-determined criteria and controls for factors that influence HRV data were imposed in data analysis. Firstly, affective data as reported by the participants were used to assign participants episode to either a positive affect or a negative affect condition. Matched timed HR data from these episodes was then extracted and categorised by episode. Each participants' 24 HRV data were grouped in to 288 x 5-minute segments, with 4 days per participant, resulting in a sample pool of 1152 x 5-minute epochs per participant (it should be noted that owing to battery capacity constraints of the Actiheart device, and given the required settings for the current study, a number of participants only returned 3x day Actiheart data). Episodic data were then sampled, and 5-minute epochs were selected using a random number generator. Once selected these data were subjected to a checklist that included the pre-determined control measures, including, accelerometer data (with a zero-movement threshold imposed for the entire 5 minute), episode type (that excluded any social stressors), social contexts (these were limited, e.g., such as no group interaction), limits on context such as driving, or commutes were excluded. Finally, where participants report in their additional qualitative data that stressor events are present, these data are excluded from analysis. A sample checklist is available in table 6.1. For the final

analysis, only ambulatory positive affect was included for comparison to laboratory conditions. Despite the presence of a number of negative affect contexts, episodes of negative affect did not provide sufficient data for analysis once the above-mentioned controls were imposed on the data.

Sampling criteria for ambulatory affect for each selected 5- minute epoch

Table 6.1: *A sample of checklist for affect context and ambulatory HRV*

Criteria	Report
Participant number	23
Day number	3
Randomly identified episode number	166
Time of day	19:45
Episode -social interaction	Alone
Episode-location	Home
Accelerometer data	No movement
Qualitative data	No confounds indicated
Heart rate data quality	Above threshold
Heart rate data	Segment 3
Randomly chosen 5 min segment	No interruptions

Data as described above were captured for each participant for six positive and six negative affect epochs for each participant. Samples were distributed across the full four -day measurement period, with two positive and two negative examples, were sought for each participant and subject to these controls. HRV data that matched these timeframes including, BPM, RMSSD and high frequency components were then extracted. These data were then conducted using the following statistical methods and presented as follows.

Multilevel Modelling

Laboratory and ambulatory heart rate data were analysed, using a hierarchical multilevel model, in IBM SPSS (IBM, 2005; Shek & Ma, 2011), In order to assess patterns of reactivity and variability a repeated measures two-level model structure for experimental and semi-continuous variables (Hoffman & Rovine J., 2007) was used. Level one represents the individuals HRV measure in context. Level two represents contexts, including, laboratory baseline, laboratory stress conditions (maths task and anticipatory singing) and ambulatory affect conditions. HRV measures were analysed separately, and data were analysed for each respective measure with fixed effects for level two, which indicates if an interclass difference exists as per context. In addition, the covariance of intercept and slopes for individuals, representing participants initial values and the degree of intrapersonal change across contexts were calculated using estimates of model improvement for random intercepts and slopes were also where fixed effects were detected, and these are presented as X^2 in log likelihood using critical values for chi squared statistic with a value of $p>0.0$. As samples of less that 50 participants can lead to biased estimates of standard error (Hox et al., 2015) the Hurvich and Tsai's criterion, suitable for small sample sizes, was employed to asses log likelihood value (Field, 2013).

6.4 Findings

6.4.1 Overview

The following describes the mean differences between test conditions for time domain and frequency components of HR. Repeated measures analysis of variance were conducted to compare of BPM, IBI, RMSSD and HF HRV during ambulatory positive affect condition, baseline laboratory, laboratory reactivity conditions including maths stress task and an anticipatory maths task. Results for each measure are detailed including self-report momentary affect scales. In addition to analysis of variance, multilevel modelling was conducted, and the results are presented.

6.4.2. Affective patterns

Table 6.2 shows the mean ambulatory affect for all participants. Ambulatory data are grouped by activity for the four-day collection period. The results show that overall positive affect have higher mean scores (Mean= 3.25, range = 2.56 – 4.00) than negative affect (Mean 0.59, range = 0.30 – 1.12). The highest positive mean affect scores were reported for Intimate relations (N=9), Socialising (N=48), and Family time (N=20). The highest Negative affect scores were returned for Work (N=109), Commute (N=144), and Study (N=9).

Table 6.2: *Overall Mean ambulatory affect rating per episode type*

Mean ambulatory affect rating per episode type														
Affect description	Happy	Competent	Warm/ friendly	Enjoying Myself	Overall Positive	Impatient	Frustrated/ Annoyed	Depressed/ Blue	Hassled/ Pushed	Angry/ Hostile	Worried/ Anxious	Criticised/ Put down	Tired	Overall Negative
Mean score Activity	3.32	3.38	3.28	3.04	**3.25**	1.30	0.86	0.36	0.27	0.24	1.00	0.10	2.13	**0.59**
Waking	3.09	3.41	3.11	2.73	**3.09**	1.26	1.04	0.54	0.50	0.46	0.99	0.13	2.96	**0.70**
Commute	2.84	3.26	2.80	2.31	**2.80**	2.32	1.25	0.56	0.44	0.43	1.36	0.21	2.23	**0.94**
Work	2.66	3.49	3.03	2.18	**2.84**	2.89	1.55	0.48	0.63	0.43	1.14	0.11	1.91	**1.03**
Socialising	3.95	3.54	3.95	4.06	**3.88**	0.60	0.43	0.06	0.33	0.25	0.60	0.10	1.68	**0.34**
Exercising	3.68	3.86	3.43	3.56	**3.63**	1.07	0.41	0.29	0.11	0.07	0.84	0.07	1.64	**0.41**
Relaxing	3.59	3.31	3.21	3.48	**3.40**	0.42	0.42	0.29	0.10	0.09	0.85	0.03	2.64	**0.31**
Prayer/Med	3.16	3.33	3.08	3.08	**3.16**	1.50	0.08	0.66	0.58	0.08	0.58	0.25	1.90	**0.53**
Phone/Skype	3.15	2.69	3.46	3.00	**3.08**	1.07	1.07	0.46	0.15	0.30	1.15	0.23	2.15	**0.63**
Internet	2.87	3.29	2.77	2.45	**2.85**	1.48	1.38	0.54	0.25	0.03	1.29	0.00	2.35	**0.71**
Watching TV	3.47	3.32	3.16	3.60	**3.39**	0.53	0.78	0.50	0.04	0.47	1.23	0.14	2.00	**0.53**
Housework	3.29	3.58	3.21	2.78	**3.22**	2.04	0.97	0.46	0.21	0.21	0.58	0.17	1.39	**0.66**
Hobby	3.40	3.50	3.62	3.81	**3.58**	0.72	0.81	0.14	0.00	0.22	0.83	0.00	0.61	**0.39**
Family time	4.00	3.80	3.80	3.65	**3.81**	0.4	1.25	0.20	0.05	0.05	0.75	0.00	1.7	**0.39**
Intimate Rel	4.11	3.77	4.11	4.00	**4.00**	0.55	0.55	0.66	0.44	0.22	0.88	0.33	1.22	**0.52**
Self-care	3.16	3.23	2.80	2.56	**2.94**	1.43	0.86	0.53	0.26	0.23	1.46	0.03	2.13	**0.69**
Eating	3.62	3.50	3.43	3.36	**3.48**	0.46	0.44	0.27	0.18	0.16	0.78	0.09	1.49	**0.34**
Driving	2.75	2.75	2.75	2.00	**2.56**	1.75	1.00	0.00	0.00	0.00	1.75	0.00	2.75	**0.64**
Shopping	3.70	3.65	3.6	3.4	**3.59**	1.35	0.65	0.10	0.30	0.15	0.65	0.00	1.75	**0.46**
Study	3.27	3.45	3.63	2.63	**3.25**	2.18	0.81	0.45	0.27	0.09	1.81	0.00	3.27	**0.80**
Bedtime	3.31	3.00	3.15	3.26	**3.18**	0.63	0.42	0.31	0.00	0.31	0.31	0.15	3.15	**0.30**
Other	2.71	3.33	2.71	2.04	**2.70**	2.66	2.09	0.14	0.76	0.85	1.19	0.14	1.61	**1.12**

Table 6.3: *Mean affect rating by laboratory condition*

	Laboratory affect ratings					
	Positive affect			Negative affect		
Condition						
	Mean	sd	*p*	Mean	sd	*p*
Baseline	4.54	0.51		1.82	1.74	
Maths task	2.90	0.97	*<0.01**	2.62	0.94	*=0.18*
Anticipatory singing	2.48	1.01	*<0.01**	3.19	1.12	*<0.01**
Recovery	4.09	1.05	*=0.32*	1.89	1.08	*=1.00*

*Indicates statistically significant difference from baseline

Table 6.3 reports laboratory affective data are grouped by their respective condition, baseline, maths stressor, anticipatory singing, and recovery. Affective scores show higher positive affect for baseline condition (Mean= 4.54, sd=0.51), and lower negative affect mean (Mean = 1.82, sd= 1.74

Positive affect

A one-way repeated measures ANOVA was conducted to compare positive affect across Baseline, Maths stress task, Sing-A-Song anticipation, and recovery. There was a significant overall effect for positive affect, Wilks' Lambda, 0.15, F (3,23) = 43.28, p < .000, multivariate partial eta squared = 0.85. Post hoc pairwise comparisons using a Bonferroni correction showed differences from baseline (M=4.538, sd =0.51) and (i) maths condition for positive affect (Mean difference = -1.64, p < .000 (ii) Sing a song condition IBI (Mean difference = -2.06, p < .000). Differences were also detected between recovery IBI (M=4.09, sd=1.05) and (i) maths condition (Mean difference = -1.92, p < .000; (ii) Sing a song condition (Mean difference = -1.615, p < .000). No statistically significant differences were found for positive affect between the two stressor conditions, or between baseline and recovery conditions. The greater magnitude mean difference between both baseline and recovery was between those conditions and the anticipatory singing task.

Negative affect

A one-way repeated measures ANOVA was also conducted to compare negative affect across Baseline, Maths stress task, Sing-A-Song anticipation, and Recovery. There was a significant overall effect for negative affect, Wilks' Lambda, 0.39, F (3,23) = 11.95, $p < .000$, multivariate partial eta squared = 0.61. Post hoc pairwise comparisons using a Bonferroni correction showed affective differences from baseline for Sing a song condition alone (Mean difference = -1.34, $p < .01$). Differences were also detected between recovery (M=4.09, SD=1.05) and Sing a song condition (Mean difference = -0.73, $p < .01$). No statistically significant differences were found for negative affect between the two stressor conditions, or between baseline and recovery conditions. The greater magnitude mean difference between both baseline and recovery was between those conditions and the anticipatory singing task.

Table 6.4. *Analysis of Variance for Resting, Reactivity, Recovery Laboratory and Ambulatory Positive affect conditions - Tests of within subjects effects*

Measure	**Model**				**Baseline**	**Maths**	**Singing**	**Recovery**	**Amb Positive**
	F	***df***	***P***	***Partial eta²***	Mean SD	Mean SD	Mean SD	Mean SD	Mean SD
*BPM **	22.337	4	0.00*	.540	67.05 10.50	77.85 11.41	79.05 13.63	71.90 10.17	69.51 8.50
Range					51 - 83	57 - 101	59 - 115	58 - 88	52 - 84
*RMSSD**	3.59**	4	0.40*	.159					
Range					41.69 19.85	30.15 11.41	32.03 5.64	51.72 41.26	50.07 27.02
					12.28 – 93.49	14.08 – 73.04	10.85 – 75.21	14.67 - 200.73	12.27 – 107.52
HF	0.429	4	0.78	.022					
Range					477.70 424.36	475.6 688.74	405.75 419.21	530.26 973.70	624.48 583.17
					29 - 1778	37 - 3092	25 – 2204	22 - 4551	32 - 1828

**Significant at the 0.05 level*

***Mauchly's test of sphericity violated – Greenhouse geisser value reported*

Beats per minute

A one-way repeated measures ANOVA was conducted to compare BPM during Baseline, Maths stress task, Sing-A-Song anticipation, and mean scores for ambulatory positive and ambulatory positive affect contexts. There was a significant overall effect for BPM, Wilks' Lambda = .16, F (4, 20) = 22.337, p < .00, multivariate partial eta squared = .540. Post hoc tests using a Bonferroni correction showed differences from baseline BPM (Mean= 66.05, SD=10.50) and (i) maths condition for BPM (Mean difference = 10.80, p < .00; (ii) Sing a song condition BPM (Mean difference = 12.00, p <.00) (iii) Recovery (Mean difference = 4.85, p < .00). No statistically significant differences were found for BPM between the two laboratory stressor conditions. No statistically significant differences were detected between Baseline laboratory BPM and ambulatory positive affect BPM

Root Mean Sum of Successive Differences

A one-way repeated measures ANOVA was conducted to compare RMSSD during Baseline, Maths stress task, Sing-A-Song anticipation, and ambulatory positive affect. There was a significant effect for RMSSD, Maulchys test indicated that sphericity had been violated so degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity, F (1.90, 20) = 3.59, p < .04, multivariate partial eta squared = .159. Post hoc tests using a Bonferroni correction showed the only pairwise differences for RMSSD were between laboratory maths task (Mean= 30.15, SD=11.41) and ambulatory positive affect (Mean difference = -19.24, p = .039. No other statistically significant differences were detected.

High frequency

A one-way repeated measures ANOVA was conducted to compare the HF component of HR during Baseline, Maths stress task, Sing-a-Song anticipation, and ambulatory positive affect conditions. There was no overall significant effect for HF. Post hoc analysis did not reveal any pair wise mean differences between conditions.

6.4.3 HRV Measures - Multilevel modelling

Heart rate variability measures of BPM, RMSSD and HF components were used to assess patterns of reactivity and variability utilising multilevel modelling with a repeated measures two-level model structure. Level one represented the individuals HRV measure per context; level two represented the separate contexts of, laboratory baseline, laboratory stress conditions (maths task and anticipatory singing) and ambulatory positive affect. Results are presented and detailed by HRV measure and respective context. Fixed effects for level two context, which indicate if an interclass difference per group exists are presented and details of significance are indicated where relevant. In addition, covariance patterns, of intercept and slope, representing participants initial values and the degree agreement of intrapersonal change across contexts is also presented.

Table 6.5: *Multilevel model for BPM, RMSSD and High frequency measures of HRV in laboratory and ambulatory contexts*

	BPM		***RMSSD***		***HF***	
Condition	Mean	*sd*	Mean	*sd*	Mean	*sd*
Baseline	67.05	*10.50*	41.69	*19.85*	447.70	*424.36*
Math task	77.85	*11.41*	30.15	*11.41*	475.60	*688.74*
Singing	79.05	*13.63*	32.03	*5.64*	405.75	*419.21*
Recovery	71.90	*10.17*	51.72	*41.26*	530.26	*973.70*
Ambulatory T1	68.10	*9.82*	42.71	*21.89*	682.90	*879.84*
Ambulatory T2	71.67	*11.12*	40.85	*26.74*	479.19	*618.05*
Ambulatory T3	69.19	*9.82*	48.18	*47.84*	488.66	*660.88*
Ambulatory T4	72.33	*11.08*	53.32	*44.93*	1000.76	*1437.62*
Ambulatory T5	69.06	*11.51*	54.26	*39.75*	827.74	*993.47*
Ambulatory T6	67.50	*10.75*	74.17	*117.80*	581.56	*730.40*

	Estimates of Covariance – Repeated measures			*Estimates of fixed effects*				
BPM	*B*	*se*$_b$	*p*	*B*	*se*$_b$	*f*	*p*	*95% CI*
Laboratory Reactivity	***30.19***	**6.67**	<0.00**	6.64	1.12	30.74	<0.00*	-4.2, 9.04
Baseline x Ambulatory	**66.65**	**8.66**	<0.00	0.17	0.49	0.12	0.73	-0.82, 1.17
Reactivity x Ambulatory	**60.10**	**7.12**	<0.00**	-1.47	0.28	28.27	<0.00*	-2.01, -0.92
RMSSD								
Laboratory Reactivity	**226.16**	**44.90**	<0.00**	-4.89	2.27	4.63	0.04*	-9.4, -0.37
Baseline x Ambulatory	**2012.22**	**255.57**	<0.00	3.76	2.22	2.86	0.09	-0.65, 8.18
Reactivity x Ambulatory	**1724.42**	**197.09**	<0.00**	4.27	1.71	6.22	0.02*	+0.86, 7.68
HF								
Laboratory Reactivity	**226541.91**	**154598.95**	<0.00	-35.02	71.83	0.23	0.62	-177.94, 107.90
Baseline x Ambulatory	**554632.77**	**431596.77**	<0.00	25.79	39.88	0.42	0.52	-53.77, 105.36
Reactivity x Ambulatory	**687788.74**	**79902.41**	<0.00	-4194	34.84	1.45	0.23	-111.50, 27.62

* Indicates a statistically significant fixed effect, level two, between contexts

** Indicates a statistically significant change in estimates of covariance of intercept and slope .

Where these changes are indicated the improved beta is reported.

Beats per minute

Baseline and laboratory reactivity patterns were examined, and BPM significantly predicted reactivity to the combined laboratory stressor maths task and anticipatory singing task. Estimates of fixed effects show significant variance F (1, 61.39) = 30.74, p< .001 across conditions. The relationship between baseline and laboratory reactivity also showed significant covariance in intercepts and slopes, $X^{2\ change} = 58.41$, Cov ($u_{oj,}$, u_{1j}) =, *2*4.74, p<0.01, indicating that while individual participant baseline BPM values differed, as indicated by the intercepts, the slopes due to conditional change also significantly covaried and improved the model prediction. In addition, patterns of laboratory baseline and ambulatory positive contexts were examined, and these did not show a fixed effect for BPM across contexts, F (1, 49.37) = 0.12, p=0.73. Finally, ambulatory positive affect BPM also significantly predicts reactivity to the combined stressor conditions of maths task and anticipatory singing task, in the laboratory. Estimates of fixed effects show significant variance F (1, 151.87) = 28.27, p< .001 across conditions. The relationship between baseline and laboratory reactivity also showed significant covariance in intercepts and slopes, $X^{2\ change} = 80.20$, Cov ($u_{oj,}$, u_{1j}) = 103.48, p <0.01, indicating that while participant ambulatory positive BPM values differed as indicated by the intercepts, the slopes due to conditional change also significantly covaried and improved the model prediction.

Root mean sum of successive square differences

Baseline RMSSD significantly predicts reactivity to the laboratory stressor maths task and anticipatory singing task. Estimates of fixed effects show significant variance F (1, 81.47) = 4.64, p=0.03 across conditions. The relationship between baseline and laboratory reactivity also showed significant covariance in intercepts and slopes, $X^{2\ change} = 10.05$, Cov (u_{oj}, u_{1j}) = 226.16, p <0 .00, indicating that while participant baseline laboratory RMSSD values differed as indicated by the intercepts, the slopes due to conditional change also significantly covaried and improved the model prediction. RMSSD measures for laboratory baseline and ambulatory contexts did not show a fixed effect across contexts, F (1, 85.35) = 2.87, p=0.09. However ambulatory positive affect RMSSD did predict reactivity to the stressor maths task and anticipatory singing task, in the laboratory. Estimates of fixed effects show significant variance F (1, 73.95) = 6.22, p =0.02 across conditions. The relationship between baseline and laboratory reactivity also showed significant covariance in intercepts and slopes, $X^{2\ change} = 90.87$, Cov (u_{oj}, u_{1j}) = 172442, p <0.00 indicating that while participant ambulatory positive RMSSD values differed as indicated by the intercepts, the slopes due to conditional change also significantly covaried and improved the model prediction.

High Frequency component

HF component of HRV did not show a fixed effect across baseline and stressor reactivity contexts, F (1, 85.54) = 0.23, p =0.62. In addition, the HF component of HRV did not show a fixed effect across baseline laboratory and ambulatory positive affect, F (1, 68.87) = 0.48, p =0.52. Finally, ambulatory positive affect HF did not significantly predict reactivity to the stressor maths task and anticipatory singing task, in the laboratory with estimates of fixed effects, F (1, 66.01) = 1.448, p = 0.23 across conditions.

6.5 Discussion

This study examined patterns of cardiovascular activity in both laboratory and ambulatory contexts, using HRV measures and retrospective diary-type measures. HRV patterns were assessed and compared between laboratory stress reactivity and ambulatory conditions. The hypothesis that baseline, and reactivity patterns in the laboratory would be related to ambulatory HRV given adequate consideration of ambulatory conditions is supported, however these patterns are not observable in all conditions and measures. Our hypothesis extended to include a prediction that laboratory baseline and laboratory reactivity conditions would differ significantly, that matched positive ambulatory patterns would not differ significantly from laboratory baseline and that laboratory reactivity and ambulatory resting condition would differ significantly. The analysis presented inter (group) differences and intra (individual) patterns of HRV for the variety of conditions, and these hypotheses are discussed in regard to each respective HRV measure.

Baseline laboratory BPM significantly predicted reactivity to the combined stressor maths task and anticipatory singing task. Estimates of fixed effects were significant across conditions and the relationship between baseline and laboratory reactivity also showed significant covariance in intercepts and slope, which, when included, improved the model estimation significantly. This demonstrates that although participant inter individual baseline BPM values differed, they also significantly covaried from baseline across the stressor conditions. RMSSD followed a similar pattern with model estimates between baseline and stressor conditions in the laboratory showing a significant fixed effect, and the model was improved by the inclusion of covariance parameters for slopes and intercepts, again indicating again that while participants returned different baseline values, they also significantly covaried from baseline across stressor conditions. However, the HF component did not show a fixed effect across baseline and

stressor reactivity contexts, even after the inclusion of random intercepts and slopes. Therefore, hypothesis one is supported with regards to measures of BPM and RMSSD for this sample, both in terms of fixed effects and how the participants co-vary in initial measurement, and how they vary as they transition across context.

The findings in terms of fixed effects, for BPM and RMSSD are in agreement, with those of Jump & Dockray (2020) in a study that examined laboratory stressor reactivity to this task, although they note that no fixed effect was detected for HF. This does not align with the findings for the HF component however it should be noted that the inclusion/exclusion criteria, imposed to maintain the rigor of the ambulatory analysis, resulted in a lower sample size for data analysis. While the reduction in numbers meant differences for BPM and RMSSD were still detectable, the high frequency component analysis resulted in a much wider range of values with larger standard deviations, which is not unusual for HF measure that are sensitive to change in context and confounding factors (Fred Shaffer & Ginsberg, 2017). Therefore, while participants intercept at baseline and the covariance of the slopes were significant for BPM and RMSSD, and showed a fixed effect for 2/3 measures, it is concluded that for this study hypothesis one is supported, yet there is a limitation that regarding the inclusion of HF component here.

The second hypothesis was that when affective ratings were matched for context for laboratory baseline and positive ambulatory patterns, the two conditions would not differ significantly in terms of HRV. BPM was examined for laboratory baseline and positive affect ambulatory contexts, and these did not show a fixed effect across contexts. Similarly, RMSSD measures did not show a fixed effect across laboratory and ambulatory positive affect conditions. Finally, HF component, did not show a fixed effect across baseline laboratory and ambulatory

positive affect conditional change. While this is not a conclusion that non-significant results confirm the hypothesis, it is noted that that the participants did not differ significantly from baseline and ambulatory contexts. While this may be an indication that there is no association between the two conditions it is encouraging that no differences were detected in baseline conditions between both contexts.

The final hypothesis was that laboratory reactivity and ambulatory resting conditions would differ significantly for the range of HRV measures. BPM for ambulatory positive affect predicted laboratory reactivity to the combined stressor conditions of maths task and anticipatory singing task, in the laboratory with significant estimates of fixed effects and significant covariance in the intercepts for ambulatory positive, change in slopes and laboratory stressor conditions. Ambulatory RMSSD also predicted reactivity to the stressor conditions, in the laboratory and also had significant covariance in the intercepts for ambulatory positive, and degree of change in slopes and laboratory stressor conditions. with. The HF component did not significantly predict reactivity to the stressor maths task and anticipatory singing task, in the laboratory. Hypothesis three is supported with regards to measures of BPM and RMSSD, both in terms of the fixed effect of condition and how the participants co-vary as they move from one condition to the next. However, following similar patterns to hypothesis one with regards to the HF component, no fixed effect was detected between conditions. Therefore, it is concluded that for this study hypothesis three is supported, although again there is a limitation regarding the inclusion of the HF component. In a review of methodological issues surrounding short term HRV measures Heathers (Heathers, 2014) suggests that in instances where there are discrepancies between multiple measures (and it should be noted that they cite RMSSD and HF indices specifically) this can be resolved by appealing to the underlying physiological phenomena that each is related to. For example, RMSSD reflects beat-to-beat variance in HR

and targets vagally mediated changes reflected in HRV (Fred Shaffer & Ginsberg, 2017). The HF component, while also indicative of parasympathetic function, is influenced by respiratory function and therefore more susceptible to experimental or confounding influence. Although it could be argued that that the differences reported here between these measures could be attributed in this way, the authors suggest that further study is required to definitively support that conclusion. Heathers suggests a range of extraneous variables, for example respiration, that could contribute to these findings and these could be examined in future study that targets the precise influence of these confounds between measures in laboratory and ambulatory contexts.

Previous studies, linking laboratory and ambulatory HR patterns physiological measure included global of 24 hr patterns measures (Gerin, Rosofsky, Pieper, & Pickering, 1993;Kamarck, Schwartz, Janicki, Shiffman, & Raynor, 2003), often measured at pre-selected intervals throughout the day (Thomas. Kamarack & William R. Lovallo, 2003), or used stressor task in ambulatory contexts (Wetherell et al., 2017), such as skydiving (Dikecligil & Mujica-Parodi, 2010). These stressors may not represent typical or frequently occurring stressors (Schwerdtfeger, Schienle, Leutgeb, & Rathner, 2014), and the advantages of the controlled conditions of the research laboratory usually prevent participants from some behavioural responses, including removing themselves from the situation, engaging in health behaviour (e.g. smoking) or other responses that in turn influence cardiovascular activity. Data from this study indicate that patterns of reactive laboratory BPM and RMSSD differ from baseline and ambulatory positive affect contexts, when confounds such as movement are accounted for.

This study carried a number of limitations, including data missingness or loss owing to incomplete reporting of DRM data or loss of signal form the ECG device. The smaller sample size may explain why the results are not in accordance with previous reports (Jump & Dockray,

2020) especially with regard to the HF component. Further study, including that with sufficient datapoints to examine HF is required, as RMSSD has been linked to HF (Shaffer & Ginsberg, 2017), and both have been linked to vagal tone (Porges, 2007), the two measures were expected to show similar patterns in the present study. Finally, there was an absences of negative affect state data from the reported ambulatory data. Studies examining the DRM (Diener & Tay, 2014; Dockray et al., 2010), demonstrate consistently lower reporting of negative affect, and whether this is due to under reporting, or to what extent people actually experience negative affect that compares in intensity to that reported during laboratory stressor tasks, remains unknown. This could be addressed in future study that specifically targets negative affect affect states specifically, or by a larger sampling period for each individual that could present more epochs for comparison. The central aim of this study was to examine the concordance of measures of heart rate variability taken in laboratory and ambulatory contexts and the findings have indicated there is some initial evidence for an alignment of laboratory HRV baseline and reactivity profiles. A final recommendation is that further research should examine the association of affect and psychobiological state across laboratory and ambulatory contexts, this will both inform evaluations of the validity of models that test associations of cardiovascular reactivity and psychobiological wellbeing using laboratory research designs.

Chapter 7: Discussion

7.1 General Discussion

The PhD presented here describes a series of studies designed to examine the accordance of measures obtained in laboratory settings and measures obtained in ecological setting. The research uses a psychobiological framework and includes measures of heart rate variability and experience sampling approaches. A review of ambulatory measurement challenges and three studies concerned with laboratory and ambulatory measurement, that employ physiological measurement techniques, were presented. Study One, "Measuring the psychobiological correlates of daily experience in adolescents" (Dockray, O'Neill & Jump, 2019) outlines the application of ambulatory psychophysiological measurement concerning a specific cohort, namely adolescents. Study Two, "Cardiovascular Responses to anticipatory stress utilising anticipatory singing tasks" (Jump & Dockray, 2020), examines an adaptation of a stressor paradigm proposed by Brouwer, (Brouwer & Hogervorst, 2014) and integrates technical adjustments to the methodology making it suitable for HRV analysis. Study Three, "Examining the use of an online adaptation of the Day Reconstruction Method", presents an adaptation of the Day Reconstruction Method (DRM) originally proposed by Kahneman and colleagues (Kahneman, Krueger, Schkade, et al., 2004). Study Four, "From the lab to the living room: Measuring heart rate variability in ecologically valid contexts", details HRV measurement in ambulatory contexts, laboratory-based measures, ambulatory HRV, and experience sampling. These measures are used to assess the relationship of the participants' laboratory and ambulatory responses.

Whilst each study may be considered individually, they combine to allow examination of general themes of psychophysiological measurements in controlled and uncontrolled settings, and the following discussion examines thematic aspects, rather than specific study findings.

This approach provides a framework that contextualises each study in relationship to the others, enabling the related theoretical underpinnings to be described. These themes include, laboratory measurement of HRV and affect, capturing experience and affect inside and outside of the laboratory, matching contexts using affect and experience, the relationship of affect to laboratory and ambulatory HRV measures, and also final conclusions.

7.2 Laboratory measurement of HRV and affect

Laboratory measurement of HRV and affect were included in Studies Two and Four. In Study Two a novel stressor procedure was piloted and in Study Four it was employed as the mode of comparison to ambulatory measures. The aim of Study Two was to assess if tasks that use an anticipatory singing prompt are a viable method for inducing stress in laboratory contexts, while adhering to a design suitable for studies using HRV measurement. The original procedure was redesigned to align with recommendations for experimental planning, data analysis and reporting for HRV type studies employing a Resting/Reactivity/Recovery structure detailed by Laborde et al., (2017 and Quintana & Heathers, (2014). Analysis from Study Two indicated that the stressor tasks utilising maths tasks and singing as stimuli were effective in generating a stress response. Statistically significant differences in HR measures between baseline in both stressor conditions for BPM, IBI, RMSSD and HF components were detected, although a number of caveats exist in relation to reactivity patterns between baseline and the respective stress tasks, particularly when individual HRV measures were considered.

Patterns of HRV measures for each task component showed that while there were overall differences between stressor conditions and recovery, only BPM and IBI were statistically significant for both stressor tasks. RMSSD and HF were not significantly different between baseline and the maths task, or maths task and recovery. These findings were unexpected; in fact it was anticipated that the maths task, which combines a social evaluative threat with the

additional cognitive load of the arithmetic task would result in higher magnitude responses. Previous research examining the association between cognitive tasks and CVR responses utilising the TSST (Allen et al., 2014, 2017) report greater magnitude responses with maths tasks plus speaking, and it was expected that similar patterns, particularly given the removal of the physiological confounds of standing and speaking. In fact, the singing task alone elicited the higher response.

One explanation for these findings is that the anticipatory singing task carries an isolated and particular social evaluative threat that results in higher magnitude responses than the arithmetic task, and this is detectable in HRV measures that can differentiate social context, namely RMSSD and HF. The relationship between social functioning, HF components of HRV, and RMSSD has been widely documented and both of these measures have been associated (Shahrestani et al., 2015) with social evaluative threat. Decreased patterns in HF and RMSSD across tasks has been offered as an explanation of social evaluative threat via vagal withdrawal (Porges, 2007; Smith et al., 2017) and increased reactivity, indexed by BPM, is resultant of increased metabolic demand, and reactivity to performance based tasks (Wetherell et al., 2017, 2014). The findings in the studies here suggest that the combination of performance task (maths) and the social evaluative threat (singing) presented here can be employed to further examine the intricacies of responses related to cognitively driven tasks and those with a social evaluative threat. In addition to allowing the task elements described here in combination with traditional TSST (Kudielka et al., 2007) procedures, the short epochs are of appropriate duration to be used with procedures that examine performance task, e.g. multitasking framework.

The next phase of the study here examined how laboratory evoked HRV patterns related to participants ambulatory data. The aim was to examine baseline and reactive states in

ambulatory and laboratory contexts, using the self-reported DRM and affective data as a means to anchor HRV data in both contexts. In study four multilevel modelling was used to analyse the laboratory and ambulatory data, owing to the complexity of the data and the nesting of epochs in ambulatory conditions (Field., 2013). Estimates of fixed effects were significant across conditions for BPM and the relationship between baseline and laboratory reactivity also showed significant covariance in intercepts and slope, which, when included, improved the model estimation significantly, demonstrating that participant inter-individual baseline BPM values differed and significantly covaried from baseline. RMSSD also differed between baseline and stressor conditions in the laboratory with significant fixed effects. Improvements to the model by the inclusion of covariance parameters for slopes and intercepts indicated again that while participants returned different baseline values, they also significantly covaried from baseline across the stressor conditions.

However, in a divergence from the findings in Study Two the HF component of HRV did not show a fixed effect across baseline and stressor reactivity contexts, even after the inclusion of random intercepts and slopes. It should be noted that the HF component almost achieved significance, and the slightly smaller sample size could be the contributory factor in these differences.

In study four strict inclusion/exclusion criteria were imposed to maintain the rigor of the ambulatory analysis, and this resulted in a lower sample size for both the laboratory and ambulatory data analysis. While the reduction in numbers did not change the outcome for BPM and RMSSD it did for HF component. As the HF component analysis usually results in a much wider range of values (Shaffer & Ginsberg, 2017) with larger standard deviations, statistically significant results are more difficult to achieve. Given that measures of BPM and RMSSD were significant, and HF components, although not significant trended that way, this allowed for

their inclusion for analysis of ambulatory contexts. To conclude, for Study Four, as with study two the hypothesis that the stressor procedure can elicit a stress response is supported, and while the previous study demonstrates BPM, RMSSD and HF measures differ significantly, there was a limitation regarding the inclusion of the HF component in Study Four and further study is required to examine these differences. This may be suggestive of a general limitation of ambulatory capture of HRV measures, where RMSSD and BPM can be used in combination, but the HF component may be problematic. However as mentioned further study with a larger sample size might achieve significance, and this study as highlighted the caution with which data using frequency components of HRV in ambulatory contexts should be interpreted.

Finally, while the studies presented here have found significant associations between laboratory and ambulatory measures of HRV there are a number of limitations. While the Actiheart device represents a workable solution for collection of ECG data and HRV extrapolation, the study procedure did reveal some limitations, namely the access to proprietary algorithm, in particular where this relates to data sorting and cleaning. In fact, this highlighted a pressing methodological concern for any future study concerned with examining HRV in ambulatory contexts. The exacting standards of measurement that are implemented in the laboratory are there to monitor participants and ensure the quality of the data. The study procedure attempted to account for this by allowing only data with a perfect 'quality' score and zero movement to be included in the analysis. Proponents of exact HRV measurement will balk at the analysis of this type of data, and their concerns are not unwarranted. It is true that the further study diverges from the lab more difficult it is to ensure its quality, however, there is a point at which counting every single heartbeat and manually examining confounds will make the task of ambulatory measurement impossible. This however is not to exclude their use entirely and what has been highlighted is the considerations that are required for research that will venture out of the lab.

7.3 Capturing experience and affect inside and outside of the laboratory

The transferability of laboratory obtained measures of stress and heart rate variability to ecological, non-laboratory settings is the foundation stone of most models of health psychology. There have been some studies to determine the strength of accord between the measures obtained in the laboratory and in an ecologically obtained measures (Konjarski et al., 2018) few attempt to match periods of similar emotional content and duration. Doing so is an essential requirement to elaborate on the relationship of emotions heart rate and behavior (Loeffler & Peper, 2014). One of the challenges is eliciting a range of emotions and this is a partial explanation of why stress is so often used in studies of psychophysiology (Allen et al., 2014), it that is reasonably easy to elicit feelings of stress and much more difficult to listen feelings of joy contentment and even more so to partner those with specific activities that are only likely to occur in ecological contexts. It is a challenging area of study to in part to develop laboratory stress challenges that truly represent the challenges that represent the stressors people may encounter in ecological settings

The study presented here posited that experience and affect, along with controls afforded by the measurement device, e.g., accelerometer data, would be sufficient to define these contexts and match them to the laboratory measures. Therefore, Study Four captured self-report affective responses across the entire laboratory procedure and ambulatory contexts, and Study Two also captured laboratory affect for all conditions. Participants were required to report affect at six intervals in the laboratory, pre-baseline, baseline, maths task, singing task, and recovery. The affective data revealed significant overall effects for across conditions in the laboratory with differences from baseline for the maths condition, singing condition, and recovery phase. The participants also indicated marginally higher, but significant, self-report

stress during the recovery period from baseline.

The ability to map affective patterns from the laboratory to ambulatory contexts, along with contextualising data from the DRM and the sensor data from the HR device, represented one of the key challenges in the study program. The affective data were used to indicate context, and the reporting of these data, in the absence of any physiological confounds, were used to match conditions inside and outside of the laboratory. However, as with the sing a song procedure, a number of methodological concerns needed to be addressed before the inclusion of the affective data from the DRM instrumentation. Central to this was participant adherence to collection protocols, as this was required to ensure the accurate description of context for the extraction of HRV data. One of challenges identified during the design phase of the studies was the apparent lack of data available to track adherence using the DRM, and it was sought to address this by amending the procedure. The procedural adjustments and the examination of these questions are presented in study three and the advantages are described.

The ability to quantify adherence represents a major advantage for the use of the DRM and in Study Three adherence patterns were demonstrated using electronic modes of collection. Although previous studies had used online collection methods, for example Dockray and colleagues (Dockray et al., 2010) describe online entry of DRM data, they did not report adherence patterns specifically. Similarly, Daly et al (2010) describe the use of electronic surveys to collect DRM data (Daly et al., 2010) and describe how the format aids participation and analysis; however, this is not further elaborated upon to describe adherence patterns. Study three described participant adherence patterns and provided novel insight into how participants report data using retrospective type measures. For example, time stamped data were compared to the timepoints participants were instructed to complete data entry and this resulted in a three-stage categorisation of adherence, full adherence, partial adherence, and non-completion, that

was used in the study program. Full adherence was described as the completion of all DRM measures, within the specified timeframe, for all four days. Partial completion was defined as full completion of data entry with deviation from the specified timeframe. The ability to confidently include or exclude DRM data represents a novel finding for DRM type studies and allowed the inclusion of the experiential and affective data for comparison with the laboratory data in study four.

The patterns observed here also raised a number of interesting questions in relation to the use of retrospective type measures. Firstly, with the use of pen and paper formats or other online methods without time stamps, are existing methods of capturing adherence patterns in the DRM reliable? Although the findings here demonstrated that non-adherence to measures is relatively low, the inability of previous methods to discriminate between participants who adhered to collection protocols and those that didn't is problematic. However, the level of granularity provided by the timestamped data in our protocol allows data entry to be screened for non-compliance. This also means that researchers can discriminate between partial and fully adherent, which allows the inclusion or exclusion of data for individual participants, where they previously would have needed to be excluded entirely, or perhaps more worrying for the researcher, they may not have known on which days the participant were compliant. This was evident in Study Three by the relatively frequent number of partial entries and grouping of data entry. In the context of the current PhD where biological data were being collected this was also particularly useful.

The second issue raised in relation to adherence is whether participants who are non-adherent or partially adherent with pen and paper formats may understand that there is no oversight once they leave the laboratory and engage in ambulatory collection phases. This may be a

contributory factor to partial adherence, as participants may prioritise adherence over time precision, for example, do they complete the measures late rather than not at all. In addition to measuring adherence, the described methods could negate this partial adherence by informing the participants that their data are time stamped and entry can be observed by the researchers in real time. The protocol presented here allows for real time interaction with the data and subsequently, possible real-time interaction with the participants. For example, the researcher can observe adherence across the study period, and initiate contact with the participant to query adherence, or prompt completion, if required. This type of protocol implemented at a laboratory briefing may be able to negate non-adherence and improve adherence patterns as whole , and with a better understanding of adherence and timestamped data, retrospective diary type measures can be more reliably integrated with other experience sampling measures. This is desirable where data from momentary and retrospective patterns can be interrogated together by researchers providing better insight into episodic and affective patterns.

Daily affective patterns as measured by the DRM have been detailed since the original DRM study (Kahneman, Krueger, Schkade, et al., 2004), in a study by Stone (Stone et al., 2006), and in relation to other experience sampling methodologies (Dockray et al., 2010), in larger cross cultural studies, (Mellor-Marsá et al., 2016), and in population level studies (Anusic et al., 2017). The collection method employed here was shown to be a robust means of accessing affect across different samples and the data collected for the four-day measurement period were similar to those described by Kahneman (2004) in the original DRM study, a sample using random sampling of weighted episodes described by Anusic (2017) and the comparison study of the DRM and ESM by Dockray (2010). Two studies by Kahneman and Anusic were selected and affective data was examined by correlating it with our sample for type of episode and social interaction and found that there were good correlations between the data from the existing

studies and study three despite applying a different statistical technique related to weighting. The implication for the collected studies presented here is that in light of the adaptation of the collection method the affective data follow similar patterns to established studies and could confidently track adherence to DRM data entry.

In addition to adherence in the application of the DRM here it was considered that there are number of statistical approaches to analysing and using this type of data, in particular the affective data that are returned by the collection method. Diener & Tay (2014) highlight the need for further work to understand the degree to which these analytical approaches influence the outcomes, with issues such as time weighting of episodes, and grouping of affect scores to generate net affect requiring further consideration. In Kahneman and colleagues original study (Kahneman, Krueger, Schkade, et al., 2004) affective ratings are not time weighted and the overall affect ratings are considered at face value, although there is a suggestion that further study could contextualise the findings in relation to episode type and social interaction, although no indication is given to the purpose of this further investigation. A subsequent study (Stone et al., 2006) again using analyses that do not time weight the data report similar findings and supports the original DRM. In a comparative study between EMA and DRM, Dockray and colleagues (2010) elaborated on the idea of time weighting and employed statistical techniques to the affective scores that were used to compare DRM and ESM methodologies. This study (Dockray et al., 2010) reported better correlations between affect for the two measures if ratings were attenuated for time spent on task. Kim et al, (2013) addressed a similar comparison between the two measurement types but did not employ weighting in their analysis, instead opting to average across episodes, similar the study presented here, the original DRM, and the Stone study. However, while Kim does report correlations between EMA and DRM measures

they note that these correlations are weak, similar to Dockray, et al, where the data are not attenuated for time.

The study presented here detailed affective patterns and the findings indicated the relationships remain robust despite the non-weighting of episodes. The decision not to employ weighted means was taken due to the small sample size and predominance of positive affect. The underlying assumption in studies that attempt to attenuate for the effect of time on affect is that time spent on task is representative of the influence of the affect associated with that particular episode, i.e., that positive affect states will carry motivation or incentive for people to remain in that state for longer. Other authors (Lee et al., 2017) go further and posit that even more complex models are required, where statistical analysis need to account for activity type, time of day, social interaction, and friendliness of interacting partners and suggests that the net affect scores generated by these models are more representative of daily experience. The question of which approach best represents retrospective human experience by applying varying statistical methods is not answered here in this thesis, although, given the similarity of the affective scores returned it does not preclude any data generated by our collection method from theses analysis later. In fact, it better facilitates it with more robust data – such as that described for adherence- that can be integrated if required.

7.4 Matching laboratory and ambulatory contexts using affect and experience

Examining if and how a relationship between laboratory and ambulatory affective and biological states exists was a goal of this research, and the focus of Study Four. Matching these data depended on participant reports of positive and negative affect states and it was posited that affective states inside the laboratory mirrored those in everyday contexts. Examining the affective and experiential data across contexts contributes to the reliability and precision of data used to anchor HR measures in experience, by starting with the individuals' own reflection

on their context in the diary, and then examining the corresponding physiological data, which itself had strictly controlled inclusion and exclusion criteria. However, as this relied on the participants reports of these data as the starting point, it was sought to ensure that these patterns were reliable and reflective of previous study.

The patterns of positive and negative affect in the study here align with other studies that used the DRM. For example, positive affect is reported more often than negative affect (Kahneman, Krueger, Schkade, et al., 2004). Kahneman describes reports of negative affect as "relatively infrequent". Kahneman also reported fluctuating patterns of affect, with more pronounced peaks and troughs across the day than the findings in the study here, although Kahneman and colleagues (2004) only report on negative affect and tiredness.

The daily patterns in the study here follow more closely those described by Anusic and colleagues (Anusic et al., 2017), where positive affect dominates and negative affect is reported to a much lesser degree. Similar to Anusic, the study here returned lower positive affect scores also matching data from larger European studies (Mellor-Marsá et al., 2016; Möwisch et al., 2019). The data presented here, although taken from a much smaller sample size, match the diurnal patterns of these European samples more closely. Similarly patterns of tiredness followed a characteristic diurnal pattern with the highest reported scores in the early morning and evening, with a small peak during late afternoon. Again, the diurnal patterns of affect observed in the study here more closely align with the data reported in the European and indicated much flatter affective patterns throughout the day than Kahnemans observations, with similar agreement regarding patterns of tiredness. These observations may indicate cultural differences, or simply be a measurement artifact of the different studies. The observed patterns of affect and tiredness in here matched existing studies and allowed confidence in the

robustness of the collection method, even given that the method had been adapted. However, there are a number of caveats in relation to the reporting of negative affect and these had an impact on the ability to report physiological data for these epochs

The design for Study Four included measures to collect positive and negative affective states. The original study design aimed to use six positive affect and six negative affect ambulatory epochs to extract HR data and then run a comparison with laboratory conditions of stress by matching negative affect with stress and positive affect with baseline. However, given the notably low reporting of negative affect enough "negative" epochs were not retrieved to run a comparison to laboratory stressor conditions, therefore only positive affect states were included for comparison with laboratory conditions. The infrequency of negative affect reports may be considered to indicate a measurement issue, a response bias, or perhaps just a very happy participant group! It may be important for future research to consider sample size, sampling frequency and sampling duration in order to generate a sufficiency of negative affect states to compare to laboratory affect, rather than use a different measurement.

It is not possible for this current work to answer the question of why positive affect is reported more frequently than negative affect. Possible explanations include recall bias inherent in retrospective type measures (Diener & Tay, 2014) and examining these data in relation to momentary reports may answer this question. It could also be related to an unwillingness to report negative affect. Another possibility is that the majority of people might actually just be happy most of the time! Incredible as this conclusion might be, it is an important note as it could give valuable insight into the assumptions about how laboratory stressors translate to real world experiences. Just how typical and frequently occurring are these stressor we examined states that we seek to replicate so carefully in the lab? Ultimately if these data are to have utility

for the individual it will be dependent to how they interpret their own data as they relate to their own experience, and ultimately their physiology. We will have to examine this question further by tracking affect and experience in controlled and ecological settings, with a range of stressors and uplifts and the experiences of daily life and this has implications for the biological data that we capture across these states.

7.5 The relationship of affect to laboratory and ambulatory HRV measures

Study Four examined patterns of cardiovascular activity in both the laboratory and ambulatory contexts to examine the concordance of the HRV measures when examined by affect type. The findings indicated that HRV patterns from laboratory and ambulatory contexts could be associated. The data were able to predict reactivity in the laboratory based on both ambulatory resting contexts and baseline measures in the laboratory. Resting positive affect conditions and laboratory baselines predicted reactivity in both BPM and RMSSD reactivity and these measures also had significant covariance. This finding represents the most novel and potentially useful finding in the thesis presented here. To our knowledge the use of affective ratings and carefully defined inclusion/exclusion criteria have not been used to match laboratory and ambulatory measures in this way. It should be noted that the HF component did not significantly predict ambulatory resting and reactivity to the stressor task in the laboratory.

Previous research has linked laboratory and ambulatory HR patterns, however these have not had the specificity to context described in Study Four. For example, most commonly in health related studies (Gerin, Rosofsky, Pieper, & Pickering, 1993; Kamarck, Schwartz, Janicki, Shiffman, & Raynor, 2003) global or 24 hr patterns measures are used to study individual patterns of HRV. While these approaches have been reliably associated with health outcomes, the study presented here contributes to a more granular understanding of the

association between laboratory and real-world contexts. If sampling is to be representative and targeted to context, approaches such as those used here will have better utility because it will allow definition of targeted context in real world settings and allow comparison to standardised laboratory measures. It is suggested here that this more targeted analysis, via HRV data extracted from carefully described contexts, will better facilitate the transition from laboratory to ambulatory measurement.

Other studies have capitalised on events and situations that are known or assumed to be stressful and occur in ecological contexts, including , for example, a skydiving event (Dikecligil & Mujica-Parodi, 2010; Wetherell et al., 2017) , driving tasks (Healey & Picard, 2005), stressful work environments such as call centres (Hernandez et al., 2011), and firefighters (Schwerdtfeger & Dick, 2019). However, these stressors may not represent typical or frequently occurring stressors for the individual (Schwerdtfeger, Schienle, Leutgeb, & Rathner, 2014). The limitation of study to these specific contexts has most likely been owing to methodological limitations; for example Smets (2019) describes the challenges associated with collection of physiological data from longer timeframes, including, physical activity and signal quality from devices, the reliance on subject defined stress, and finally the difficulty in defining baseline conditions in ambulatory contexts.

It is possible to resolve these challenges and it has been demonstrated here that the use of appropriate experience sampling, HR sensor data, and the careful description of context via strict inclusion/exclusion criteria, allow for the association of laboratory and ambulatory contexts. This represents an attempt to link laboratory and ambulatory context by sampling randomly throughout the day and then including the data if-and-as they meet the sampling criteria. The ease of this approach will be enabled by advances in technologies to collect

measures with low burden and high precision, however as yet there is no consensus on the adoption of devices and application of algorithms, and this is owing to the lack of evidence of precision, validity, and reliability (Alexander, Minhaj Uddin & Joshi, 2017; Plante et al., 2016), of current collection methods. Therefore, conclusions about the utility of this type of research, such as the methodology proposed here, must be made with caution, but are a promising avenue for how ambulatory psychophysiological processes in ecological contexts will be matched to standardised measures. The current research demonstrates how these data can be tentatively linked by examining the individuals reports of their own experience and mood, and then by overlaying strict, pre-defined HRV measurement criteria. The current findings demonstrate that, for the reported HRV measures, these contexts can be linked and this provides preliminary but promising evidence for further work to scaleup, using larger data sets, in particular those that are seeking criteria by which to define measurement periods for HRV extraction from ambulatory contexts. This could also have applications for the development of HRV measures for research or commercial application using momentary experience sampling, for example, to co-relate physiological data with psychological and environmental context data for use in health applications, or for just-in-time technologically mediated interventions.

For researchers the findings also indicate the possibilities of conducting pre-laboratory measures, for example the procedures and data analysis methods can potentially be used for pre-laboratory attendance measures to augment traditional laboratory procedures or inform how the reactivity and recovery profiles might be anticipated by the individual-specific typical patterns of response generated for each person. For example, a profile of baseline affect could be collected in ambulatory contexts that could be potentially used in conjunction with previous methods as a basis of comparison to laboratory measures, and to account for confounds such as white coat effects (Leventhal et al., 2007). If health psychology is to ultimately provide the

individual with access to their own data and allow them to interact with it to improve health, then anchoring it in experience with which they are familiar should be the end goal.

7.6 Conclusions

Since the Task Force Paper (Camm et al., 1996) on HRV research several dominant research threads have emerged, including, biomedical research focused on the technical requirements of HRV measurement, medical research concerned with disease aetiology and clinical applications (Sassi et al., 2015), and psychophysiological sciences that aim to describe how HRV is associated with psychosocial variables, such as affect (Balzarotti et al., 2017), psychopathologies (Beauchaine et al., 2019), and health (Chambers & Allen, 2007). The research studies described in this dissertation are most relevant to the strand concerned with psychophysiological science and contribute to knowledge of the biopsychosocial frameworks. The studies contribute to the understanding of rhythms and change in physiology and mood, across laboratory and ambulatory contexts, and add to knowledge about experience sampling, indicating how technology can reduce the burden as well as increase precision of measure of daily mood and experience. In addition, the individual contributions sit within the dominant theoretical frameworks that see to match social context and physiological response to environmental cues.

Thayer's (Thayer et al., 2012) and Porges (Porges, 2001) theories emerged as the dominant frameworks and provided an explanation of how individuals navigate complex social interactions that relate to HRV indices. The underlying principle was that HRV could be used as a marker of top down and affective regulation and this has been widely supported (Holzman & Bridgett 2017). In Polyvagal Theory, Porges posited that individuals self-regulate based on environmental cues and that lowered HRV is indicative of parasympathetic suppression

(Porges et al, 2001, 2007). The levels of activation detailed in NVI describe how environmental context and demands will determine the level of autonomic activation, and subsequently how top-down regulation of these responses can be indexed using HRV measures. For example Thayer and colleagues suggest that HRV measures can be used to differentiate between elements of cognition (e.g., working memory), affective responses, autonomic regulation (Thayer et al., 2009). Study Two presented two such differential tasks and the findings indicate that the responses detected in HRV measures were related to the specificity of each task, in particular the different social and cognitive challenges. A stressor paradigm is described that examined resting, reactive and recovery affect and HRV patterns in a laboratory context. This study demonstrated the use of a novel social stimulus that adhered to measurement standards for HRV analysis and examined the social context of this type of measurement. Tonic and phasic HRV are described in relation to an anticipatory type stressor and their relevance for the specificity to social context. However as stated previously the examination of these theories had primarily been carried out in laboratory contexts (Laborde et al., 2017) and the study here extended the comparison of HRV patterns in the subsequent ambulatory study.

The dominant research paradigms to date have contributed significantly to our understanding of HRV measures and their psychosocial correlates. (Smets et al., 2019), and Porges and Thayer's work is heavily focused on the social context of humans. Therefore, capturing everyday social context with situational, non-invasive measurement of HRV may enable better predictions of how prolonged exposure to stressors result in HRV measurements that are linked to health (Smith, Deits-Lebehn, Williams, Baucom, & Uchino, 2020). Thayer specifies how neurology and function differentiate based on experience and how that experience interacts with the processing of current events, and this has implications for the relevance of the data captured in laboratory, for example, if participants are exposed to novel or infrequently

occurring stressors, then how representative is it of their reaction to everyday events? Moreover, if behavioural mediation has an important role in people's exposure to stressors (McEwen, 2001), situationally located measurement will be better able to describe how people navigate everyday environments and may be more reflective of how they interact with or avoid stressors. How people navigate, mediate, and interpret the stressors they encounter (or don't as the low reports of negative affect in the current data suggested) is just as relevant as the measurement of the effect of those stressors on individuals' physiology, and this relates to how experience is integrated into their psychology. Therefore, the interpretation of experience must resolve the issue of integrating momentary experience and retrospective data.

Despite the most sophisticated collection methods available , the question of whether the experiencing self or the remembered (Zajchowski et al., 2017) self is the best method of interrogating experience remains. Although this may be a false dichotomy, and experience sampling, if it is to be truly reflective of real-world psychological phenomena and affect, will need to integrate the two. As Kahneman demonstrated using a fictious pain stimulus, participants accounts will vary substantially from momentary data, particularly where that experience is negative (Redelmeier & Kahneman, 1996a). Some experiences may be perceived as negative or stressful at the time, and retrospective accounts may tend to reduce negative recall. In addition, it is subject to the myriad biases that humans use to integrate experience once viewed in retrospect. Patterns of affective and episodic data in our study would seem to suggest that this is idea is supported, where, as mentioned, participants report the lowest number of episodes for work periods, even though the majority of their waking time is spent at work. Also, the underreporting of negative affect, which is a feature of the entire collection phase, may indicate, as Kahneman suggests, that participants are filtering or not reporting these

instances. Future study of experience will need to better understand how these patterns are experienced and remembered by the individual.

While proponents of momentary sampling may question the validity of retrospective measures, citing these issues, and this leads to suggesting real-time assessment is the most accurate and valid measure. However, the way people retrospectively interpret their experiences is a fascinating puzzle, and very salient to understanding of how the integration of experience into psychophysiological health occurs. Both the phenomenon of momentary experience and how our remembering self-integrates experiences need to be captured by our instrumentation, and these need to be measured accurately in order to account for the whole experience of our participants. The work presented here has demonstrated how the adaptation of electronic methods that time stamp and monitor data will allow more effective integration of the two methods, which will benefit both participant and researchers and ultimately capture better data. As technology advances and allows people to interpret their own data, a baseline of research is required to contextualise and authenticate the various claims, in particular as that experience relates to physiological measures.

Research that pairs ambulatory and laboratory HRV data and psychosocial measures with measures obtained in everyday life still presents considerable methodological challenges. This thesis has detailed some of those challenges, proposing a means for assessing HRV across contexts using experience sampling, affect, and physical sensor data. While there are still myriad challenges, resolving them will contribute to how these data can ultimately be of use in real world contexts. How the technologies that capture these data are integrated into peoples' experience, in non-intrusive ways in their own environments, that allow them to reflect on and integrate them into their own experience is the ultimate goal. As these technologies become

available to the majority the coal face of our science will is moving with them. If the story of heart rate measurement in humans *really is* the story of the technology available to measure it at any given time, the next chapter of that story will be about how we make the journey with them as our science moves from lab to the living room.

References

Adam, E. K., & Kumari, M. (2009). Assessing salivary cortisol in largescale, epidemiological research. Psychoneuroendocrinology, 34(10), 1423-1436.

Adam, E. K., Quinn, M. E., Tavernier, R., McQuillan, M. T., Dahlke, K. A., & Gilbert, K. E. (2017). Diurnal cortisol slopes and mental and physical health outcomes: A systematic review and meta-analysis. Psychoneuroendocrinology.

Alexander, J. C., Minhajuddin, A., & Joshi, G. P. (2017). Comparison of smartphone application -based vital sign monitors without external hardware versus those used in clinical practice: a prospective trial. Journal of clinical monitoring and computing, 31(4), 825-831

Alink, L. R., van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., Mesman, J., Juffer, F., & Koot, H. M. (2008). Cortisol and externalizing behavior in children and adolescents: Mixed meta-analytic evidence for the inverse relation of basal cortisol and cortisol reactivity with externalizing behavior. Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology, 50(5), 427-450.

Allen, A. P., Kennedy, P. J., Cryan, J. F., Dinan, T. G., & Clarke, G. (2014). Biological and psychological markers of stress in humans: Focus on the Trier Social Stress Test. In Neuroscience and Biobehavioral Reviews (Vol. 38)

Allen, A. P., Kennedy, P. J., Dockray, S., Cryan, J. F., Dinan, T. G., & Clarke, G. (2017). The Trier Social Stress Test: Principles and practice. Neurobiology of Stress, 6, 113–126.

Antelmi, I., De Paula, R. S., Shinzato, A. R., Peres, C. A., Mansur, A. J., & Grupi, C. J. (2004). Influence of age, gender, body mass index, and functional capacity on heart rate variability in a cohort of subjects without heart disease. American Journal of Cardiology, 93(3), 381–385. https://doi.org/10.1016/j.amjcard.2003.09.065

Anusic, I., Lucas, R. E., & Donnellan, M. B. (2017). The Validity of the Day Reconstruction Method in the German Socio-economic Panel Study. Social Indicators Research, 130(1), 213–232. https://doi.org/10.1007/s11205-015-1172-6

Arvidson, E., Sjörs, A., & Jonsdottir, I. H. (2017). Perceived stress and physiological reactions to repeated TSST in healthy individuals. Psychoneuroendocrinology, 83, 15.

Balzarotti, S., Biassoni, F., Colombo, B., & Ciceri, M. R. (2017). Cardiac vagal control as a marker of emotion regulation in healthy adults: A review. In Biological Psychology (Vol. 130). https://doi.org/10.1016/j.biopsycho.2017.10.008

Bandodkar, A. J., & Wang, J. (2014). Non-invasive wearable electrochemical sensors: a review. Trends in biotechnology, 32(7), 363-371.

Barold, S. S. (2003). Willem Einthoven and the birth of clinical electrocardiography a hundred years ago. Cardiac Electrophysiology Review, 7(1), 99–104.

Barreira, T., Kang, M., & Caputo, J. (2009). Validation of the Actiheart monitor for the measurement of physical activity. International Journal of Exercise Science, 2, 60–71.

Bartlett, J. (2014). The “Weisure” Class: The Elusive Goal of Work-Life Balance. Library Leadership & Management, 28(2).

Behar, J., Roebuck, A., Domingos, J. S., Gederi, E., & Clifford, G. D. (2013). A review of current sleep screening applications for smartphones. Physiological measurement, 34(7), 29.

Beltzer, E. K., Fortunato, C. K., Guaderrama, M. M., Peckins, M. K., Garramone, B. M., & Granger, D. A. (2010). Salivary flow and alpha-amylase: collection technique, duration, and oral fluid type. Physiology & behavior, 101(2), 289-296.

Beauchaine, T. P., Bell, Z., Knapton, E., McDonough-Caplan, H., Shader, T., & Zisner, A. (2019). Respiratory sinus arrhythmia reactivity across empirically based structural dimensions of psychopathology: A meta-analysis. Psychophysiology, 56(5), 1–19. https://doi.org/10.1111/psyp.13329

Benarroch, E. (1993). The central autonomic network: functional organization, dysfunction, and perspective. Mayo Clinic Proceedings, 68(10), 988–1001.

Bernardi, L., Valle, F., Coco, M., Calciati, A., & Sleight, P. (1996). Physical activity influences heart rate variability and very-low-frequency components in Holter electrocardiograms. Cardiovascular Research, 32(2), 234–237. https://doi.org/10.1016/0008-6363(96)00081-8

Bertz, J. W., Epstein, D. H., & Preston, K. L. (2018). Combining ecological momentary assessment with objective, ambulatory measures of behavior and physiology in substance-use research. Addictive Behaviors, 83(June 2017), 5–17. https://doi.org/10.1016/j.addbeh.2017.11.027

Bhat, S., Ferraris, A., Gupta, D., Mozafarian, M., DeBari, V. A., Gushway-Henry, N.,Chokroverty, S. (2015). Is There a Clinical Role For Smartphone Sleep Apps? Comparison of Sleep Cycle Detection by a Smartphone Application to Polysomnography. Journal of Clinical Sleep Medicine 11(7), 709-715.

Bianchi, M. T. (2015). Consumer Sleep Apps: When it Comes to the Big Picture, it's All About the Frame. J Clin Sleep Med, 11(7), 695-696. doi:10.5664/jcsm.4834

Billman, G. E. (2011). Heart rate variability - A historical perspective. Frontiers in Physiology, (November), 1–13. https://doi.org/10.3389/fphys.2011.00086

Billman, G. E. (2013). The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance. Frontiers in Physiology, 4 FEB(February), 1–5. https://doi.org/10.3389/fphys.2013.00026

Birkett, M. a. (2011). The Trier Social Stress Test protocol for inducing psychological stress. Journal of Visualized Experiments : JoVE, 56, 1–6. https://doi.org/10.3791/3238

Biskup, M. J., Kaplan, S., Bradley-Geist, J. C., & Membere, A. A. (2019). Just how miserable is work? A meta-analysis comparing work and non-work affect. PloS one, 14(3), e0212594.

Brage, S., Brage, N., Ekelund, U., Luan, J., Franks, P. W., Froberg, K., & Wareham, N. J. (2006). Effect of combined movement and heart rate monitor placement on physical activity estimates during treadmill locomotion and free-living. European Journal of Applied Physiology, 96(5), 517–524. https://doi.org/10.1007/s00421-005-0112-6

Brannon, E. E., Cushing, C. C., Crick, C. J., & Mitchell, T. B. (2016). The promise of wearable sensors and ecological momentary assessment measures for dynamical systems modeling in adolescents: A feasibility and acceptability study. Translational behavioral medicine, 6(4), 558 -565.

Brouwer, A. M., & Hogervorst, M. A. (2014). A new paradigm to induce mental stress: The Sing-a-Song Stress Test (SSST). Frontiers in Neuroscience, 8(8 JUL), 1–8. https://doi.org/10.3389/fnins.2014.00224

Brown, L., Karmakar, C., Gray, R., Jindal, R., Lim, T., & Bryant, C. (2018). Heart rate variability alterations in late life depression: A meta-analysis. In Journal of Affective Disorders (Vol. 235). https://doi.org/10.1016/j.jad.2018.04.071

Brown, S. B. R. E., Brosschot, J. F., Versluis, A., Thayer, J. F., & Verkuil, B. (2017). New methods to optimally detect episodes of non-metabolic heart rate variability reduction as an indicator of psychological stress in everyday life. International Journal of Psychophysiology. https://doi.org/https://doi.org/10.1016/j.ijpsycho.2017.10.007

Burke, H. M., Davis, M. C., Otte, C., & Mohr, D. C. (2005). Depression and cortisol responses to psychological stress: A meta-analysis. Psychoneuroendocrinology, 30(9), 846–856. https://doi.org/10.1016/j.psyneuen.2005.02.010

Burke, L. E., Shiffman, S., Music, E., Styn, M. A., Kriska, A., Smailagic, A., Siewiorek, D., Ewing, L. J., Chasens, E., French, B., Mancino, J., Mendez, D., Strollo, P., & Rathbun, S. L. (2017). Ecological momentary assessment in behavioural research: Addressing technological and human participant challenges. Journal of Medical Internet Research, 19(3). https://doi.org/10.2196/jmir.7138

Byrne, E. A., & Porges, S. W. (1993). Data-dependent filter characteristics of peak-valley respiratory sinus arrhythmia estimation: A cautionary note. Psychophysiology, 30(4), 397–404.

Câmara, R., & Griessenauer, C. J. (2015). Chapter 27 - Anatomy of the Vagus Nerve (R. S. Tubbs, E. Rizk, M. M. Shoja, M. Loukas, N. Barbaro, & R. J. B. T.-N. and N. I. Spinner (eds.); pp. 385–397). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-410390-0.00028-7

Camm, A. J., Malik, M., Bigger, J. T., Breithardt, G., Cerutti, S., Cohen, R. J., Coumel, P., Fallen, E. L., Kennedy, H. L., & Kleiger, R. E. (1996). Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology.

CamNtech. (2010a). The Actiheart Document Reference : Health (San Francisco), 35(Jan), 1–95.

CamNtech. (2010b). The Actiheart Guide to Getting Started. Options, 37(Apr), 1–39. http://www.camntech.com/files/The_Actiheart_Guide_to_Getting_Started.pdf

CamNtech. (2018). Actiheart user manual. https://doi.org/10.15713/ins.mmj.3

Capon, H., Hall, W., Fry, C., & Carter, A. (2016). Realising the technological promise of smartphones in addiction research and treatment: an ethical review. International Journal of Drug Policy, 36, 47 -57.

Carissimi, A., Martins, A. C., Dresch, F., da Silva, L. C., Zeni, C. P., & Hidalgo, M. P. (2016). School start time influences melatonin and cortisol levels in children and adolescents–a community-based study. Chronobiology international, 33(10), 1400-1409.

Carskadon, M. A., Vieira, C., & Acebo, C. (1993). Association between puberty and delayed phase preference. Sleep, 16(3), 258-262.

Centre Substance Abuse Treatment. (1994). Simple Screening Instrument for Substance Abuse Figure H-1 . Sources for Items Included in the Simple Screening Instrument for Substance Abuse. 1–13.

Chalmers, J. A., Quintana, D. S., Abbott, M. J. A., & Kemp, A. H. (2014). Anxiety disorders are associated with reduced heart rate variability: A meta-analysis. Frontiers in Psychiatry, 5(JUL). https://doi.org/10.3389/fpsyt.2014.00080

Chambers, A. S., & Allen, J. J. B. (2007). Cardiac vagal control, emotion, psychopathology, and health. Biological Psychology, 74(2), 113–115. https://doi.org/10.1016/j.biopsycho.2006.09.004

Chen, Z., Lin, M., Chen, F., Lane, N. D., Cardone, G., Wang, R., ... & Campbell, A. T. (2013, May). Unobtrusive sleep monitoring using smartphones. In Pervasive Computing Technologies for Healthcare (Pervasive Health), 2013 7th International Conference on (pp. 145 -152). IEEE.

Chen, F. R., Raine, A., & Granger, D. A. (2015). Tactics for modeling multiple salivary analyte data in relation to behavior problems: Additive, ratio, and interaction effects. Psychoneuroendocrinology, 51, 188- 200.

Cheung, M. N., & Porges, S. W. (1977). Respiratory influences on cardiac responses during attention. Physiological Psychology, 5(1), 53–57. https://doi.org/10.3758/BF03335298

Chida, Y., & Hamer, M. (2008). Chronic Psychosocial Factors and Acute Physiological Responses to Laboratory-Induced Stress in Healthy Populations: A Quantitative Review of 30 Years of Investigations. Psychological Bulletin, 134(6), 829–885. https://doi.org/10.1037/a0013342

ChuDuc, H., NguyenPhan, K., & NguyenViet, D. (2013b). A Review of Heart Rate Variability and its Applications. APCBEE Procedia, 7(Supplement C), 80–85. https://doi.org/https://doi.org/10.1016/j.apcbee.2013.08.016

Clow, A., Hucklebridge, F., & Smyth, N. (2018). Hormonal measurement in psychobiological research. The Routledge international handbook of psychobiology, 113-126.

Cohen, H. (2001). Association of autonomic dysfunction and clozapine: Heart rate variability and risk for sudden death in patients with schizophrenia on long-term psychotropic medication. The British Journal of Psychiatry, 179(2), 167–171. https://doi.org/10.1192/bjp.179.2.167

Ciolfi, L., & Lockley, E. (2018). From Work to Life and back again: Examining the digitally-mediated work/life practices of a group of knowledge workers. Computer Supported Cooperative Work (CSCW), 27(3), 803-839.

Crouter, S. E., Churilla, J. R., & Bassett, D. R. (2008). Accuracy of the Actiheart for the assessment of energy expenditure in adults. European Journal of Clinical Nutrition, 62(6), 704–711. https://doi.org/10.1038/sj.ejcn.1602766

Crowley, S. J., Acebo, C., & Carskadon, M. A. (2007). Sleep, circadian rhythms, and delayed phase in adolescence. Sleep medicine, 8(6), 602-612.

Csikszentmihalyi, M. (2006). The experience of psychopathology: Investigating mental disorders in their natural settings. Cambridge University Press.

Csikszentmihalyi, M., & Larson, R. (2014). Validity and reliability of the experience-sampling method. In Flow and the foundations of positive psychology(pp. 35–54). Springer.

Csikszentmihalyi, Mihaly. (2014). Flow and the foundations of positive psychology: The collected works of Mihaly Csikszentmihalyi. Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi, 1–298. https://doi.org/10.1007/978-94-017-9088-8

Dahl, R. E., & Lewin, D. S. (2002). Pathways to Adolescent Health: Sleep Regulation and Behavior. Journal of Adolescent Health, 31, 175-184.

Daly, M., Delaney, L., Doran, P. P., Harmon, C., & MacLachlan, M. (2010). Naturalistic monitoring of the affect-heart rate relationship: a day reconstruction study. Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association, 29(2), 186–195. https://doi.org/10.1037/a0017626

de Looff, P. C., Cornet, L. J. M., Embregts, P. J. C. M., Nijman, H. L. I., & Didden, H. C. M. (2018). Associations of sympathetic and parasympathetic activity in job stress and burnout: A systematic review. PLoS ONE, 13(10).

deVries, M. W. (1992). The experience of psychopathology: Investigating mental disorders in their natural settings.

Diener, E., & Tay, L. (2014). Review of the Day Reconstruction Method (DRM). Social Indicators Research, 116(1), 255–267. https://doi.org/10.1007/s11205-013-0279

Dikecligil, G. N., & Mujica-Parodi, L. R. (2010). Ambulatory and Challenge-Associated Heart Rate Variability Measures Predict Cardiac Responses to Real-World Acute Emotional Stress. Biological Psychiatry, 67(12), 1185–1190. https://doi.org/https://doi.org/10.1016/j.biopsych.2010.02.001

Dockray, S., O'Neill, S., & Jump, O. (2019). Measuring the psychobiological correlates of daily experience in adolescents. Journal of Research on Adolescence, 29(3), 595-612. https://doi.org/10.1111/jora.12473

Dockray, S., Stone, A. a, & Steptoe, A. (2010). A comparison of affect ratings obtained with ecological momentary assessment and the day reconstruction method. 99(2), 269–283. https://doi.org/10.1007/s11205-010-9578-7

Dockray, S., Smyth, N., Flynn, M., Kearney, I., Cusack, K., Mulally, A., & Collins, M. R. (2017). Picture yourself here: A novel method for determine sampling protocol adherence in ecological settings. Psychoneuroendocrinology, 83, 16 -17. https://doi.org/10.1016/j.psyneuen.2017.07.283

Dockray, S., & Steptoe, A. (2010). Positive affect and psychobiological processes. Neurosci Biobehav Rev, 35(1), 69-75. 10.1016/j.neubiorev.2010.01.006

Drury, R. L., Porges, S., Thayer, J., & Ginsberg, J. P. (2019). Editorial: Heart Rate Variability, Health and Well-Being: A Systems Perspective. In Frontiers in Public Health (Vol. 7). https://doi.org/10.3389/fpubh.2019.00323

Dolsen, M. R., Wyatt, J. K., & Harvey, A. G. (2018). Sleep, Circadian Rhythms, and Risk Across Health Domains in Adolescents With an Evening Circadian Preference. Journal of Clinical Child & Adolescent Psychology, 1-11.

Drake, E. C., Sladek, M. R., & Doane, L. D. (2016). Daily cortisol activity, loneliness, and coping efficacy in late adolescence: A longitudinal study of the transition to college. International Journal of Behavioral Development, 40(4), 334 -345.

Dunton, G. F., Liao, Y., Intille, S. S., Spruijt-Metz, D., & Pentz, M. (2011). Investigating Children's Physical Activity and Sedentary Behavior Using Ecological Momentary Assessment With Mobile Phones. Obesity, 19(6), 1205-1212.

Duval, S., Fujisawa, K., & Hashizume, H. (2007, November). Fundamental needs in wearable computing: Specificities for young people. In Convergence Information Technology, 2007. International Conference on (pp. 1326 -1334). IEEE.

Ellis, R. J., Zhu, B., Koenig, J., Thayer, J. F., & Wang, Y. (2015). A careful look at ECG sampling frequency and R- peak interpolation on short-term measures of heart rate variability. Physiological Measurement, 36(9).

Esco, M. R., & Flatt, A. A. (2014). Ultra-short-term heart rate variability indexes at rest and post-exercise in athletes: evaluating the agreement with accepted recommendations. Journal of Sports Science & Medicine, 13(3), 535.

Field, A. (2013). Discovering statistics using IBM SPSS statistics. Sage.

Fishel, S. R., Muth, E. R., & Hoover, A. W. (2007). Establishing Appropriate Physiological Baseline Procedures for Real-Time Physiological Measurement. Journal of Cognitive Engineering and Decision Making, 1(3), 286–308. https://doi.org/10.1518/155534307X255636

Floyer, S. J. (1710). The Pulse Watch. Heart rate variability – A historical perspective George E. Billman 2011

Forbes, E. J., & Porges, S. W. (1973). Heart-rate classical conditioning with a noxious auditory stimulus in human newborns. Psychophysiology, 10(2), 192–193.

Fracasso, M. P., Porges, S. W., Lamb, M. E., & Rosenberg, A. A. (1994). Cardiac activity in infancy: Reliability and stability of individual differences. Infant Behavior and Development, 17(3), 277–284.

Fuligni, A. J., & Hardway, C. (2006). Daily variation in adolescents' sleep, activities, and psychological well- being. Journal of Research on Adolescence, 16(3), 353-378.

Fuligni, A. J., Arruda, E. H., Krull, J. L., & Gonzales, N. A. (2018). Adolescent sleep duration, variability, and peak levels of achievement and mental health. Child development, 89(2), e18-e28.

Gamble, A. L., D'Rozario, A. L., Bartlett, D. J., Williams, S., Bin, Y. S., Grunstein, R. R., & Marshall, N. S. (2014). Adolescent Sleep Patterns and Night-Time Technology Use: Results of the Australian Broadcasting Corporation's Big Sleep Survey. PLoS One, 9(11), e111700.

Garcia, C., Hardeman, R. R., Kwon, G., Lando-King, E., Zhang, L., Genis, T., . . . Kinder, E. (2014). Teenagers and Texting: Use of a Youth Ecological Momentary Assessment System in Trajectory Health Research With Latina Adolescents. JMIR mHealth and uHealth, 2(1), e3.

Gary G. Berntson, Karen S. Quigly, & D. L. (2007). Cardiovascular Psychophysiology. In Handbook of Psychophysiology (pp. 182–222).

Geisler, F. C. M., Kubiak, T., Siewert, K., & Weber, H. (2013). Cardiac vagal tone is associated with social engagement and self-regulation. Biological Psychology, 93(2).

Gerin, W., Rosofsky, M., Pieper, C., & Pickering, T. G. (1993). A test of reproducibility of blood pressure and heart rate variability using a controlled ambulatory procedure. Journal of Hypertension, 11(10), 1127—1131. http://europepmc.org/abstract/MED/8258678

Gerra, G., Zaimovic, A., Zambelli, U., Timpano, M., Reali, N., Bernasconi, S., & Brambilla, F. (2000). Neuroendocrine responses to psychological stress in adolescents with anxiety disorder. 42(2), 82–92.

Goldschmidt, A. B., Wonderlich, S. A., Crosby, R. D., Engel, S. G., Lavender, J. M., Peterson, C. B., Mitchell, J. E. (2014). Ecological momentary assessment of stressful events and negative affect in bulimia nervosa. J Consult Clin Psychol, 82(1), 30-39.

Granger, D. A., Kivlighan, K. T., Fortunato, C., Harmon, A. G., Hibel, L. C., Schwartz, E. B., & Whembolua, G. L. (2007). Integration of salivary biomarkers into developmental and behaviorally-oriented research: problems and solutions for collecting specimens. Physiology & behavior, 92(4), 583-590.

Greenspan, S. I., & Porges, S. W. (2018). Psychopathology in Infancy and Early Childhood : Clinical Perspectives on the Organization of Sensory and Affective-Thematic Experience Wiley on behalf of the Society for Research i. 55(1), 49–70.

Gröschl, M. (2017). Saliva: a reliable sample matrix in bioanalytics. Bioanalysis, 9(8), 655-668.

Hagenauer, M. H., Perryman, J. I., Lee, T. M., & Carskadon, M. A. (2009). Adolescent changes in the homeostatic and circadian regulation of sleep. Developmental neuroscience, 31(4), 276-284.

Guzik, P., Piekos, C., Pierog, O., Fenech, N., Krauze, T., Piskorski, J., & Wykretowicz, A. (2018). Classic electrocardiogram-based and mobile technology derived approaches to heart rate variability are not equivalent. *International journal of cardiology*, *258*, 154-156.

Hang, A., Von Zezschwitz, E., De Luca, A., & Hussmann, H. (2012, October). Too much information! user attitudes towards smartphone sharing. In Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design (pp. 284-287).

Hare, O. A., Wetherell, M. A., & Smith, M. A. (2013). State anxiety and cortisol reactivity to skydiving in novice versus experienced skydivers. Physiology and Behavior, 118, 40–44. https://doi.org/10.1016/j.physbeh.2013.05.011

Hatch, J. P., Klatt, K., Porges, S. W., Schroeder-Jasheway, L., & Supik, J. D. (1986). The relation between rhythmic cardiovascular variability and reactivity to orthostatic, cognitive, and cold pressor stress. Psychophysiology, 23(1), 48–56.

Healey, J. A., & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. IEEE Transactions on Intelligent Transportation Systems, 6(2), 156–166. https://doi.org/10.1109/TITS.2005.848368

Heathers, J. A. (2013). Smartphone-enabled pulse rate variability: an alternative methodology for the collection of heart rate variability in psychophysiological research. International Journal of Psychophysiology, 89(3), 297 -304.

Heathers, J. A. ., & Goodwin, M. S. (2017). Dead Science in Live Psychology: A Case Study from Heart Rate Variability (HRV). 1–34. https://doi.org/10.17605/OSF.IO/637YM

Heathers, J. A. J. (2014). Everything Hertz: Methodological issues in short-term frequency-domain HRV. In Frontiers in Physiology: Vol. 5 MAY. https://doi.org/10.3389/fphys.2014.00177

Hernandez, J., Morris, R. R., & Picard, R. W. (2011). with Person-Specific Models. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 6974 LNCS(PART 1), 125–134.

Hernando, D., Roca, S., Sancho, J., Alesanco, Á., & Bailón, R. (2018). Validation of the apple watch for heart rate variability measurements during relax and mental stress in healthy subjects. *Sensors*, *18*(8), 2619.

Heron, K. E., Everhart, R. S., McHale, S. M., & Smyth, J. M. (2017). Systematic Review: Using mobile-technology-based Ecological Momentary Assessment (EMA) methods with youth: a systematic review and recommendations. Journal of Pediatric Psychology, 42(May), 1087–1107. https://doi.org/10.1093/jpepsy/jsx078

Hoffman, L., & Rovine J., M. (2007). Multilevel models for the experimental psychologist: Foundations and. Behavior Research Methods, 39(1), 101–117. https://doi.org/10.3758/BF03192848

Hofmann, S. G., Moscovitch, D. A., & Kim, H. J. (2006). Autonomic correlates of social anxiety and embarrassment in shy and non-shy individuals. International Journal of Psychophysiology, 61(2), 134–142. https://doi.org/10.1016/j.ijpsycho.2005.09.003

Holter, N. J. (1961). New method for heart rate studies continuous electrocardiography of active subjects. Science 134, 1214–1220.

Holzman, J. B., & Bridgett, D. J. (2017). Heart rate variability indices as bio-markers of top-down self-regulatory mechanisms: A meta-analytic review. Neuroscience and Biobehavioral Reviews. https://doi.org/10.1016/j.neubiorev.2016.12.032

Hox, J., Maas, C. J. M., & Hox, J. J. (2015). Sufficient Sample Sizes for Multilevel Modeling European Journal of Research Methods for the Behavioural and Social Sciences, 1(3), 86–92. https://doi.org/10.1027/1614-1881.1.3.86

Huynh, V. W., Guan, S. S. A., Almeida, D. M., McCreath, H., & Fuligni, A. J. (2016). Everyday discrimination and diurnal cortisol during adolescence. Hormones and behaviour , 80, 76-81.

IBM. (2005). Linear Mixed-Effects Modeling in SPSS : An Introduction to the Mixed Procedure. Technical Report, 28.

Jackowska, M., Dockray, S., Hendrickx, H., & Steptoe, A. (2011). Psychosocial Factors and Sleep Efficiency: Discrepancies Between Subjective and Objective Evaluations of Sleep. Psychosomatic Medicine, 73(9), 810-816.

Jump, O., & Dockray, S. (2020). Cardiovascular responses to stress utilizing anticipatory singing tasks. Journal of Psychophysiology https://doi.org/10.1027/0269-8803/a000269

Juster, R. P., & Lupien, S. (2012). A sex- and gender-based analysis of allostatic load and physical complaints. Gend Med, 9(6), 511-523. doi:10.1016/j.genm.2012.10.008

Kahneman, D., Krueger, A. A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. Science, 306(5702), 1776–1780. https://doi.org/10.1126/science.1103572

Kahneman, D., Krueger, A. B., Schwarz, N., & Stone, A. (2004). The Day Reconstruction Method (DRM): Instrument Documentation. Science, 306(July), 1–56. https://doi.org/10.1126/science.1103572

Kahneman, D., & Riis, J. (2005). Living, and Thinking about It: Two Perspectives on Life. The Science of Well-Being.

Kamarck, T. W., Schwartz, J. E., Janicki, D. L., Shiffman, S., & Raynor, D. A. (2003). Correspondence between laboratory and ambulatory measures of cardiovascular reactivity : A multilevel modeling approach. 40, 675–683.

Kamarack & Lovallo. (2003). Cardiovascular Reactivity to Psychological Challenge : Conceptual and. 21, 9–21. https://doi.org/10.1097/01.PSY.0000030390.34416.3E

Kaufmann, T., Sütterlin, S., Schulz, S. M., & Vögele, C. (2011). ARTiiFACT: A tool for heart rate artifact processing and heart rate variability analysis. Behavior Research Methods, 43(4), 1161–1170. https://doi.org/10.3758/s13428-011-0107-7

Kemp, A. H., Quintana, D. S., Gray, M. A., Felmingham, K. L., Brown, K., & Gatt, J. M. (2010). Impact of Depression and Antidepressant Treatment on Heart Rate Variability: A Review and Meta-Analysis. Biological Psychiatry, 67(11), 1067–1074. https://doi.org/10.1016/j.biopsych.2009.12.012

Kim, J., Kikuchi, H., & Yamamoto, Y. (2013). Systematic comparison between ecological momentary assessment and day reconstruction method for fatigue and mood states in healthy adults. British Journal of Health Psychology, 18(1), 155–167. https://doi.org/10.1111/bjhp.12000

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test'--a tool for investigating psychobiological stress responses in a laboratory setting. In Neuropsychobiology (Vol. 28, Issues 1–2, pp. 76–81). https://doi.org/119004

Kirschbaum, C., & Hellhammer, D. H. (2000). Salivary cortisol. Encyclopedia of stress, 3(379 -383).

Kligfield, P., Gettes, L. S., Bailey, J. J., Childers, R., Deal, B. J., Hancock, E. W., Van Herpen, G., Kors, J. A., Macfarlane, P., Mirvis, D. M., Pahlm, O., Rautaharju, P., & 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 Cli. Circulation, 115(10), 1306–1324. https://doi.org/10.1161/CIRCULATIONAHA.106.180200

Koenig, J., & Thayer, J. F. (2016b). Sex differences in healthy human heart rate variability: A meta-analysis. In Neuroscience and Biobehavioral Reviews (Vol. 64). https://doi.org/10.1016/j.neubiorev.2016.03.007

Konjarski, M., Murray, G., Lee, V. V., & Jackson, M. L. (2018). Reciprocal relationships between daily sleep and mood: A systematic review of naturalistic prospective studies. *Sleep Medicine Reviews*, *42*, 47–58. https://doi.org/10.1016/j.smrv.2018.05.005

Krantz, D. S., & Manuck, S. B. (1984). Acute psychophysiologic reactivity and risk of cardiovascular disease: A review and methodologic critique. In Psychological Bulletin (Vol. 96, Issue 3, pp. 435–464). American Psychological Association. https://doi.org/10.1037/0033-2909.96.3.435

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. Biological Psychology, 84(3), 394–421. https://doi.org/10.1016/j.biopsycho.2010.03.010

Kristiansen, J., Korshøj, M., Skotte, J. H., Jespersen, T., Søgaard, K., Mortensen, O. S., & Holtermann, A. (2011). Comparison of two systems for long-term heart rate variability monitoring in free-living conditions - a pilot study. BioMedical Engineering OnLine, 10(1), 27. https://doi.org/10.1186/1475-925X-10-27

Kroenke, K., Spitzer, R. L., Williams, J. B. W., & Löwe, B. (2009). An Ultra-Brief Screening Scale for Anxiety and Depression: The PHQ–4. Psychosomatics, 50(6), 613–621. https://doi.org/10.1016/S0033-3182(09)70864-3

Kruithof, C. J., Kooijman, M. N., van Duijn, C. M., Franco, O. H., de Jongste, J. C., Klaver, C. C., ... & Rivadeneira, F. (2014). The generation R study: biobank update 2015. European journal of epidemiology, 29(12), 911-927.

Kudielka, B. M., Broderick, J. E., & Kirschbaum, C. (2003). Compliance with saliva samplingprotocols: electronic monitoring reveals invalid cortisol daytime profiles innoncompliant subjects. Psychosom Med, 65(2), 313-319.

Kudielka, B. M., Buske-Kirschbaum, A., Hellhammer, D. H., & Kirschbaum, C. (2004). HPA axis responses to laboratory psychosocial stress in healthy elderly adults, younger adults, and children: Impact of age and gender. Psychoneuroendocrinology, 29(1), 83–98. https://doi.org/10.1016/S0306-4530(02)00146-4

Kudielka, Brigitte M, Hellhammer, D. H., & Kirschbaum, C. (2007). Ten years of research with the Trier Social Stress Test—revisited. Social Neuroscience: Integrating Biological and Psychological Explanations of Social Behavior, 56–83. https://doi.org/10.4135/9781412956253.n539

Laborde, S., Mosley, E., & Thayer, J. F. (2017). Heart Rate Variability and Cardiac Vagal Tone in Psychophysiological Research – Recommendations for Experiment Planning, Data Analysis, and Data Reporting. Frontiers in Psychology, 08(February), 1–18. https://doi.org/10.3389/fpsyg.2017.00213

Lacey, J.I., 1967. Somatic response patterning and stress: some revisions ofactivation theory. In: Appley, M.H., Trumbell, R. (Eds.), Psychological Stress. Appleton- Century-Crofts, New York, pp. 14–37.

Larsen, B. A., & Christenfeld, N. J. (2011). Cognitive distancing, cognitive restructuring, and cardiovascular recovery from stress. Biological Psychology, 86(2), 143-148.

Lauderdale, D. S., Knutson, K. L., Yan, L. L., Liu, K., & Rathouz, P. J. (2008). Self-reported and measured sleep duration: how similar are they? Epidemiology (Cambridge, Mass.), 19(6), 838-845.

Lauricella, A. R., Cingel, D. P., Blackwell, C., Wartella, E., & Conway, A. (2014). The mobile generation: Youth and adolescent ownership and use of new media. Communication Research Reports, 31(4), 357 -64.

Lee, P. H., Tse, A. C. Y., & Lee, K. Y. (2017). A new statistical model for the Day Reconstruction Method. International Journal of Methods in Psychiatric Research, 26(4). https://doi.org/10.1002/mpr.1547

Levine, J. C., Fleming, R., Piedmont, J. I., Cain, S. M., & Chen, W. J. (2016). Heart rate variability and generalized anxiety disorder during laboratory-induced worry and aversive imagery. Journal of Affective Disorders, 205. https://doi.org/10.1016/j.jad.2016.07.019

Libman, E., Fichten, C. S., Bailes, S., & Amsel, R. (2000). Sleep Questionnaire Versus Sleep Diary: Which Measure Is Better? International Journal of Rehabilitation and Health, 5(3), 205-209.

Lipponen, J. A., & Tarvainen, M. P. (2019). A robust algorithm for heart rate variability time series artefact correction using novel beat classification. Journal of Medical Engineering and Technology, 43(3), 173–181. https://doi.org/10.1080/03091902.2019.1640306

Lockley, S. W., Skene, D. J., & Arendt, J. (1999). Comparison between subjective and actigraphic measurement of sleep and sleep rhythms. J Sleep Res, 8(3), 175-183.

Loeffler, S. N., & Peper, M. (2014). Affect and memory in real life: Evaluating the utility of interactive ambulatory assessment. *Zeitschrift Für Neuropsychologie*, *25*(4), 267–278. https://doi.org/10.1024/1016-264X/a000142

Lorber, M. F. (2004). Psychophysiology of Aggression, Psychopathy, and Conduct Problems: A Meta-Analysis. Psychological Bulletin, 130(4), 531–552. https://doi-org.ucc.idm.oclc.org/10.1037/0033-2909.130.4.531

Lovato, N., & Gradisar, M. (2014). A meta-analysis and model of the relationship between sleep and depression in adolescents: recommendations for future research and clinical practice. Sleep Medicine Reviews, 18(6), 521-529.

Ludwig, C. (1847). Beitrage zur Kennt- niss des Einflusses der Respriations bewegungen auf den Blutlauf im Aortensysteme. Arch. Anat. Physiol. 13, 242–302.

Ly, J., McGrath, J. J., & Gouin, J. P. (2015). Poor sleep as a pathophysiological pathway underlying the association between stressful experiences and the diurnal cortisol profile among children and adolescents. Psychoneuroendocrinology, 57, 51-60.

Magallón-Neri, E., Kirchner-Nebot, T., Forns-Santacana, M., Calderón, C., & Planellas, I. (2016). Ecological Momentary Assessment with smartphones for measuring mental health problems in adolescents. World Journal of Psychiatry, 6(3), 303–310. http://doi.org/10.5498/wjp.v6.i3.303

Malik, M., Camm, A. J., Huikuri, H., Lombardi, F., Schmidt, G., Schwartz, P. J., & Zabel, M. (2018). Electronic gadgets and their health-related claims. *Int J Cardiol, 258*, 163-164.

Martinson, B. C., Lazovich, D., Lando, H. A., Perry, C. L., McGovern, P. G., & Boyle, R. G.(2000). Effectiveness of Monetary Incentives for Recruiting Adolescents to an Intervention Trial to Reduce Smoking. Preventive Medicine, 31(6), 706-713.

Matsumura, K., & Yamakoshi, T. (2013). iPhysioMeter: A new approach for measuring heart rate and normalized pulse volume using only a smartphone. Behavior research methods, 45(4), 1272 -1278.

Matvienko-Sikar, K., & Dockray, S. (2017). Effects of a novel positive psychological intervention on prenatal stress and well-being: A pilot randomised controlled trial. Women Birth, 30(2), e111-e118.

McEwen, B. S. (2001). From molecules to mind. Stress, individual differences, and the social environment. Annals of the New York Academy of Sciences, 935, 42–49. https://doi.org/10.1111/j.1749-6632.2001.tb03469.x

McMakin, D. L., & Alfano, C. A. (2015). Sleep and anxiety in late childhood and early adolescence. Current opinion in psychiatry, 28(6), 483.

Mellor-Marsá, B., Miret, M., Abad, F. J., Chatterji, S., Olaya, B., Tobiasz-Adamczyk, B., Koskinen, S., Leonardi, M., Haro, J. M., Ayuso-Mateos, J. L., & Caballero, F. F. (2016). Measurement Invariance of the Day Reconstruction Method: Results from the COURAGE in Europe Project. Journal of Happiness Studies, 17(5), 1769–1787. https://doi.org/10.1007/s10902-015-9669

Miocevic, O., Cole, C. R., Laughlin, M. J., Buck, R. L., Slowey, P. D., & Shirtcliff, E. A. (2017). Quantitative Lateral Flow Assays for Salivary Biomarker Assessment: A Review. Front Public Health, 5, 133.

Moeller, J., Lieb, R., Meyer, A. H., Loetscher, K. Q., Krastel, B., & Meinlschmidt, G. (2014). Improving ambulatory saliva-sampling compliance in pregnant women: a randomized controlled study. PloS One, 9(1), e86204.

Moore, M., & Meltzer, L. J. (2008). The sleepy adolescent: causes and consequences of sleepiness in teens. Paediatric Respiratory Reviews, 9(2), 114-121.

Möwisch, D., Schmiedek, F., Richter, D., & Brose, A. (2019). Capturing Affective Well-Being in Daily Life with the Day Reconstruction Method: A Refined View on Positive and Negative Affect. Journal of Happiness Studies, 20(2), 641–663. https://doi.org/10.1007/s10902-018-9965-3

Mrug, S., Tyson, A., Turan, B., & Granger, D. A. (2016). Sleep problems predict cortisol reactivity to stress in urban adolescents. Physiology & behavior, 155, 95-101.

Muslukhov, I., Boshmaf, Y., Kuo, C., Lester, J., & Beznosov, K. (2013, August). Know your enemy: the risk of unauthorized access in smartphones by insiders. In Proceedings of the 15th international conference on Human -computer interaction with mobile devices and services (pp. 271 -280). ACM.

Myin-Germeys, I., Kasanova, Z., Vaessen, T., Vachon, H., Kirtley, O., Viechtbauer, W., & Reininghaus, U. (2018). Experience sampling methodology in mental health research: new insights and technical developments. World Psychiatry, 17(2), 123-132.

Nalla, A. A., Thomsen, G., Knudsen, G. M., & Frokjaer, V. G. (2015). The effect of storage conditions on salivary cortisol concentrations using an enzyme immunoassay. Scand J Clin Lab Invest, 75(1), 92-95.

Nater, U. M., Rohleder, N., Gaab, J., Berger, S., Jud, A., Kirschbaum, C., & Ehlert, U. (2005). Human salivary alpha-amylase reactivity in a psychosocial stress paradigm. International Journal of Psychophysiology, 55(3), 333-342

Nagy, T., Van Lién, R., Willemsen, G., Proctor, G., Efting, M., Fülöp, M., ... & Bosch, J. A. (2015). A fluid response: Alpha-amylase reactions to acute laboratory stress are related to sample timing and saliva flow rate. Biological psychology, 109, 111-119.

Nater, U. M., Rohleder, N., Schlotz, W., Ehlert, U., & Kirschbaum, C. (2007). Determinants of the diurnal course of salivary alpha-amylase. Psychoneuroendocrinology, 32(4), 392-401.

Niarchou, M., Zammit, S., & Lewis, G. (2015). The Avon Longitudinal Study of Parents and Children (ALSPAC) birth cohort as a resource for studying psychopathology in childhood and adolescence: a summary of findings for depression and psychosis. Social psychiatry and psychiatric epidemiology, 50(7), 1017 - 1027.

Nicolson, N. A. (2008). Measurement of cortisol. Handbook of physiological research methods in health psychology, 1, 37 -74.

Nunan, D., Sandercock, G. R. H., & Brodie, D. A. (2010). A quantitative systematic review of normal values for short-term heart rate variability in healthy adults. PACE - Pacing and Clinical Electrophysiology, 33(11), 1407–1417. https://doi.org/10.1111/j.1540-8159.2010.02841.x

O'Brien, E. M., & Mindell, J. A. (2005). Sleep and risk-taking behavior in adolescents. Behavioral sleep medicine, 3(3), 113-133.

O'Campo, P., BeLue, R., Borenstein, H., Reed -Vance, M., Lanzi, R. G., Schafer, P., ... & Woolord, R. (2016). Issues and solutions for collecting biological specimen in longitudinal studies: Experience from the Community Child Health Network Research Network. Journal of health care for the poor and underserved, 27(1), 339 - 351

O'Doherty, K., Jaszczak, A., Hoffmann, J. N., You, H. M., Kern, D. W., Pagel, K., ... & McClintock, M. K. (2014). Survey field methods for expanded biospecimen and biomeasure collection in NSHAP Wave 2. Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 69(Suppl_2), S27-S37.

O'Neill, S., & Dockray, S. (2014, September 2014). Chronotype, Genotype, and Sleep Behaviours in Adolescents; The Effect of Technology Access on Preferred and Actual Behaviour. Paper presented at the 22nd Congress of the European Sleep Research Society, Tallinn, Estonia.

O'Neill, S., & Dockray, S. (2015). Adolescent Sleep: Contributing Factors to Poor Sleep Quality and Health Outcomes. Paper presented at the 12th Annual Psychology, Health and Medicine Conference, Belfast.

Okun, M. L., Krafty, R. T., Buysse, D. J., Monk, T. H., Reynolds, C. F., Begley, A., & Hall, M. (2010). What constitutes too long of a delay? Determining the cortisol awakening

response (CAR) using self-report and PSG-assessed wake time. Psychoneuroendocrinology, 35(3), 460-468.

Olds, T., Blunden, S., Petkov, J., & Forchino, F. (2010). The relationships between sex, age, geography and time in bed in adolescents: a meta-analysis of data from 23 countries. Sleep medicine reviews, 14(6), 371-378.

Oskis, A., Loveday, C., Hucklebridge, F., Thorn, L., & Clow, A. (2009). Diurnal patterns of salivary cortisol across the adolescent period in healthy females. Psychoneuroendocrinology, 34(3), 307-316.

Pan, J., & Tompkins, W. J. (1985). A Real-Time QRS Detection Algorithm. IEEE Transactions on Biomedical Engineering, BME-32(3), 230–236. https://doi.org/10.1109/TBME.1985.325532

Pasch, K. E., Laska, M. N., Lytle, L. A., & Moe, S. G. (2010). Adolescent sleep, risk behaviors, and depressive symptoms: are they linked? American journal of health behavior, 34(2), 237-248.

Peters, R. H., Greenbaum, P. E., Steinberg, M. L., Carter, C. R., Ortiz, M. M., Fry, B. C., & Valle, S. K. (2000). Effectiveness of screening instruments in detecting substance use disorders among prisoners. Journal of Substance Abuse Treatment, 18(4), 349–358. https://doi.org/10.1016/S0740-5472(99)00081-1

Phillips, A. C., & Hughes, B. M. (2011). Introductory paper: Cardiovascular reactivity at a crossroads: where are we now?. Biological psychology, 86(2), 95-97.

Phillips, A. C., Ginty, A. T., & Hughes, B. M. (2013). The other side of the coin: Blunted cardiovascular and cortisol reactivity are associated with negative health outcomes. International Journal of Psychophysiology, 90(1), 1–7. https://doi.org/10.1016/j.ijpsycho.2013.02.002

Phillips, M. M., Phillips, K. T., Lalonde, T. L., & Dykema, K. R. (2014). Feasibility of text messaging for ecological momentary assessment of marijuana use in college students. Psychol Assess, 26(3), 947- 957.

Peirce, J. (2019). PsychoPy (Version 3) [Software]. Open Science Tools Ltd. https://www.psychopy.org

Plante, T. B., Urrea, B., Blumenthal, R. S., MacFarlane, Z. T., Miller, E. R., Appel, L. J., & Martin, S. S. (2016). Validation of the instant blood pressure smartphone app. JAMA Internal Medicine, 176(5), 700–702. https://doi.org/10.1001/jamainternmed.2016.0157

Porges, S W. (1979). Developmental designs for infancy research. Handbook of Infant Development, 742–765.

Porges, S W. (1986). An accurate method of quantifying RSA., Psychophysiology, 23(4), 414.

Porges, Stephen W. (1973). Heart rate variability: An autonomic correlate of reaction time performance. 1(4), 270–272. https://doi.org/10.3758/BF03333367

Porges, Stephen W. (1994). Orienting in a Defensive World:Mammalian modifications of our evolutionary herritage. A Polyvagal Theory. Psychophysiology, 32, 301–318.

Porges, Stephen W. (2001). The polyvagal theory: Phylogenetic substrates of a social nervous system. International Journal of Psychophysiology, 42(2), 123–146. https://doi.org/10.1016/S0167-8760(01)00162-3

Porges, Stephen W. (2007a). A phylogenetic journey through the vague and ambiguous Xth cranial nerve: A commentary on contemporary heart rate variability research. Biological Psychology, 74(2), 301–307. https://doi.org/10.1016/j.biopsycho.2006.08.007

Porges, Stephen W. (2007b). The polyvagal perspective. Biological Psychology, 74(2), 116–143. https://doi.org/10.1016/j.biopsycho.2006.06.009

Porges, Stephen W., & Raskin, D. C. (1969). Respiratory and heart rate components of attention. Journal of Experimental Psychology, 81(3), 497–503. https://doi.org/10.1037/h0027921

Porges, Stephen W., Stamps, L. E., & Walter, G. F. (1974). Heart rate variability and newborn heart rate responses to illumination changes. Developmental Psychology, 10(4), 507–513. https://doi.org/10.1037/h0036560

Porges, Stephen W. (1974). Heart rate indices of newborn attentional responsivity. Merrill-Palmer Quarterly of Behavior and Development, 20(4), 231–254.

Porges, Stephen W. (1976). Hyperactivity, Psychopathy, and Autism., Advances in Child Development and Behavior, 11, 35.

Porges, Stephen W. (1985). Spontaneous oscillations in heart rate: Potential index of stress. In Animal stress (pp. 97–111). Springer.

Porges, Stephen W. (1986). Respiratory sinus arrhythmia: Physiological basis, quantitative methods, and clinical implications. In Cardiorespiratory and cardio somatic sychophysiology (pp. 101–115). Springer.

Porges, Stephen W. (1988). Vagal tone as an index of mental state.

Porges, Stephen W, Arnold, W. R., & Forbes, E. J. (1973). Heart Rate Variability : Developmental Psychology, 8(1), 85–92.

Porges, Stephen W, Bohrer, R., Cheung, M., Drasgow, F., McCabe, P., & Keren, G. (1980). New time series statistic for Detecting Rhythmic Co-Occurrence in the Frequency Domain. In Psychological bulletin (Vol. 88, Issue 2, pp. 580–587).

Porges, Stephen W, & Bohrer, R. E. (1990). The analysis of periodic processes in psychophysiological research.

Porges, Stephen W, & Byrne, E. A. (1992). Research methods for measurement of heart rate and respiration. Biological Psychology, 34(2–3), 93–130.

Porges, Stephen W, & Smith, K. M. (1980). Defining hyperactivity: Psychophysiological and behavioral strategies. Hyperactive Children: The Social Ecology of Identification and Treatment, 75–104.

Porges, W. (2015). Vagal Tone : A Physiological Marker of Stress Vulnerability. 90(3).

Qasem, L., Cardew, A., Wilson, A., Griffiths, I., Halsey, L. G., Shepard, E. L. C., Gleiss, A. C., & Wilson, R. (2012). Tri-axial dynamic acceleration as a proxy for animal energy expenditure; should we be summing values or calculating the vector? PLoS ONE, 7(2). https://doi.org/10.1371/journal.pone.0031187

Quintana, D. S., Alvares, G. A., & Heathers, J. A. J. (2016). Guidelines for Reporting Articles on Psychiatry and Heart rate variability (GRAPH): recommendations to advance research communication. In Translational psychiatry (Vol. 6). https://doi.org/10.1038/tp.2016.73

Quintana, Daniel S., Guastella, A. J., Outhred, T., Hickie, I. B., & Kemp, A. H. (2012). Heart rate variability is associated with emotion recognition: Direct evidence for a relationship between the autonomic nervous system and social cognition. International Journal of Psychophysiology, 86(2), 168–172. https://doi.org/10.1016/j.ijpsycho.2012.08.012

Quintana, Daniel S., & Heathers, J. A. J. (2014). Considerations in the assessment of heart rate variability in biobehavioral research. In Frontiers in Psychology (Vol. 5, Issue JUL). https://doi.org/10.3389/fpsyg.2014.00805

Quintana, Daniel S., McGregor, I. S., Guastella, A. J., Malhi, G. S., & Kemp, A. H. (2013). A Meta-Analysis on the Impact of Alcohol Dependence on Short-Term Resting-State Heart Rate Variability: Implications for Cardiovascular Risk. Alcoholism: Clinical and Experimental Research, 37(SUPPL.1), 23–29. https://doi.org/10.1111/j.1530-0277.2012.01913.x

Rao, A., Hou, P., Golnik, T., Flaherty, J., & Vu, S. (2010). Evolution of data management tools for managing self-monitoring of blood glucose results: a survey of iPhone applications. Journal of diabetes science and technology, 4(4), 949 -957.

Raugh, I. M., Chapman, H. C., Bartolomeo, L. A., Gonzalez, C., & Strauss, G. P. (2019). A comprehensive review of psychophysiological applications for ecological momentary assessment in psychiatric populations. Psychological Assessment, 31(3), 304.

Rautaharju, P. M., Park, L., Rautaharju, F. S., & Crow, R. (1998). A standardized procedure for locating and documenting ECG chest electrode positions: Consideration of the effect of breast tissue on ECG amplitudes in women. Journal of Electrocardiology, 31(1), 17–29. https://doi.org/10.1016/S0022-0736(98)90003-6

Redelmeier, D. A., & Kahneman, D. (1996). Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures. Pain, 66(1), 3–8. https://doi.org/https://doi.org/10.1016/0304-3959(96)02994-6

Reiss, A., Indlekofer, I., Schmidt, P., & Van Laerhoven, K. (2019). Deep PPG: large-scale heart rate estimation with convolutional neural networks. *Sensors*, *19*(14), 3079.

Reyes del Paso, G. A., Langewitz, W., Mulder, L. J. M., van Roon, A., & Duschek, S. (2013). The utility of low frequency heart rate variability as an index of sympathetic cardiac tone: A review with emphasis on a reanalysis of previous studies. Psychophysiology, 50(5), 477–487. https://doi.org/10.1111/psyp.12027

Richard Jennings, J., Allen, B., Gianaros, P. J., Thayer, J. F., & Manuck, S. B. (2015). Focusing neurovisceral integration: Cognition, heart rate variability, and cerebral blood flow. Psychophysiology, 52(2), 214–224. https://doi.org/10.1111/psyp.12319

Roenneberg, T., Kuehnle, T., Pramstaller, P.P., Ricken, J., Havel,M/, Guth, A., and Merrow, M. "A marker for the end of adolescence." Current Biology 14, no. 24 (2004): R1038-R1039.

Rogers, A., Caruso, C. C., & Aldrich, M. S. (1993). Reliability of Sleep Diaries for Assessment of Sleep/Wake Patterns. Nursing Research, 42(6), 368-371.

Romeo, R. D. (2017). The Metamorphosis of Adolescent Hormonal Stress Reactivity: A Focus on Animal Models. Frontiers in neuroendocrinology.

Runyan, J. D., Steenbergh, T. A., Bainbridge, C., Daugherty, D. A., Oke, L., & Fry, B. N. (2013). A Smartphone Ecological Momentary Assessment/Intervention “App” for Collecting Real-Time Data and Promoting Self-Awareness. PLoS One, 8(8), e71325.

Sadeh, A. (2011). The role and validity of actigraphy in sleep medicine: an update. Sleep Med Rev, 15(4), 259- 267.

Salvo, P., Pingitore, A., Barbini, A., & Di Francesco, F. (2017). A wearable sweat rate sensor to monitor the athletes' performance during training. Science & Sports.

Sano, A., Phillips, A. J., Yu, A. Z., McHill, A. W., Taylor, S., Jaques, N., . . . Picard, R. W. (2015). Recognizing Academic Performance, Sleep Quality, Stress Level, and Mental Health using Personality Traits, Wearable Sensors and Mobile Phones. Int Conf Wearable Implant Body Sens Netw, 2015.

Sassi, R., Cerutti, S., Lombardi, F., Malik, M., Huikuri, H. V., Peng, C. K., Schmidt, G., & Yamamoto, Y. (2015). Advances in heart rate variability signal analysis: Joint position statement by the e-Cardiology ESC Working Group and the European Heart Rhythm Association co-endorsed by the Asia Pacific Heart Rhythm Society. Europace, 17(9), 1341–1353. https://doi.org/10.1093/europace/euv015

Saxbe, D. E., Negriff, S., Susman, E. J., & Trickett, P. K. (2015). Attenuated hypothalamic–pituitary–adrenal axis functioning predicts accelerated pubertal development in girls 1 year later. Development and psychopathology, 27(3), 819-828.

Schnall, R., Okoniewski, A., Tiase, V., Low, A., Rodriguez, M., & Kaplan, S. (2013). Using Text Messaging to Assess Adolescents' Health Information Needs: An Ecological Momentary Assessment. Journal of Medical Internet Research, 15(3), e54.

Schober, M. F., Conrad, F. G., Antoun, C., Ehlen, P., Fail, S., Hupp, A. L., . . . Zhang, C. (2015). Precision and Disclosure in Text and Voice Interviews on Smartphones. PLoSOne, 10(6), e0128337.

Schommer, N. C., Hellhammer, D. H., & Kirschbaum, C. (2003). Dissociation between reactivity of the hypothalamus-pituitary-adrenal axis and the sympathetic-adrenal-medullary system to repeated psychosocial stress. Psychosomatic medicine, 65(3), 450-460.

Schroeder, E. B., Liao, D., Chambless, L. E., Prineas, R. J., Evans, G. W., & Heiss, G. (2003). Hypertension, Blood Pressure, and Heart Rate Variability: The Atherosclerosis Risk in Communities (ARIC) Study. Hypertension, 42(6), 1106–1111. https://doi.org/10.1161/01.HYP.0000100444.71069.73

Schwerdtfeger, A. R., & Dick, K. (2019). Episodes of momentary resilience in daily life are associated with HRV reductions to stressful operations in firefighters: an ambulatory assessment approach using bayesian multilevel modeling. Journal of Positive Psychology, 14(5), 593–602. https://doi.org/10.1080/17439760.2018.1497689

Schwerdtfeger, A. R., Schienle, A., Leutgeb, V., & Rathner, E.-M. (2014a). Does cardiac reactivity in the laboratory predict ambulatory heart rate? Baseline counts. Psychophysiology, 51(6), 565–572. http://10.0.4.87/psyp.12199

Schwerdtfeger, A. R., Schienle, A., Leutgeb, V., & Rathner, E. M. (2014b). Does cardiac reactivity in the laboratory predict ambulatory heart rate? Baseline counts. Psychophysiology. https://doi.org/10.1111/psyp.12199

Shaffer, Fred, & Ginsberg, J. P. (2017). An Overview of Heart Rate Variability Metrics and Norms. Frontiers in Public Health, 5(September), 1–17. https://doi.org/10.3389/fpubh.2017.00258

Shaffer, Fred, McCraty, R., & Zerr, C. L. (2014). A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability. Frontiers in Psychology, 5(September), 1–19. https://doi.org/10.3389/fpsyg.2014.01040

Shaffer, Fredric, Shearman, S., & Meehan, Z. M. (2016). The Promise of Ultra-Short-Term (UST) Heart Rate Variability Measurements. Biofeedback, 44(4), 229–233. https://doi.org/10.5298/1081-5937-44.3.09

Shahrestani, S., Stewart, E. M., Quintana, D. S., Hickie, I. B., & Guastella, A. J. (2015). Heart rate variability during adolescent and adult social interactions: A metaanalysis. Biological Psychology, 105, 43 –50.

Shaw, S.M., Caldwell, L.L., & Kleiber, D.A. (1996). Boredom, Stress and Social Control in the Daily Activities of Adolescents. Journal of Leisure Research, 28(4)

Shek, D. T. L., & Ma, C. M. S. (2011). Longitudinal Data Analyses Using Linear Mixed Models in SPSS: Concepts, Procedures and Illustrations. The Scientific World JOURNAL, 11, 42–76. https://doi.org/10.1100/tsw.2011.2

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological Momentary Assessment. Annual Review of Clinical Psychology, 4(1), 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Shimamura, a P. (2000). The role of the prefrontal cortex in dynamic filtering. Psychobiology, 28(2), 207–218.

Shirtcliff, E. A., Buck, R. L., Laughlin, M. J., Hart, T., Cole, C. R., & Slowey, P. D. (2015). Salivary cortisol results obtainable within minutes of sample collection correspond with traditional immunoassays. Clin Ther, 37(3), 505-514.

Shirtcliff, E. A., Granger, D. A., Schwartz, E., & Curran, M. J. (2001). Use of salivary biomarkers in biobehavioral research: cotton-based sample collection methods can interfere with salivary immunoassay results. Psychoneuroendocrinology, 26(2), 165-173.

Short, M. A., Arora, T., Gradisar, M., Taheri, S., & Carskadon, M. A. (2017). How many sleep diary entries are needed to reliably estimate adolescent sleep?. Sleep, 40(3), zsx006.

Silvetti, M. S., Drago, F., & Ragonese, P. (2001). Heart rate variability in healthy children and adolescents is partially related to age and gender. International Journal of Cardiology, 81(2 –3), 169 –174.

Slowey, P. D. (2015). Saliva collection devices and diagnostic platforms. In Advances in Salivary Diagnostics (pp. 33-61). Springer, Berlin, Heidelberg.

Smets, E., De Raedt, W., & Van Hoof, C. (2019). Into the Wild: The Challenges of Physiological Stress Detection in Laboratory and Ambulatory Settings. IEEE Journal of Biomedical and Health Informatics, 23(2), 463–473. https://doi.org/10.1109/JBHI.2018.2883751

Smith, R., Thayer, J. F., Khalsa, S. S., & Lane, R. D. (2017). The hierarchical basis of neurovisceral integration. In Neuroscience and Biobehavioral Reviews (Vol. 75). https://doi.org/10.1016/j.neubiorev.2017.02.003

Smith, T. W., Deits-Lebehn, C., Williams, P. G., Baucom, B. R. W., & Uchino, B. N. (2020). Toward a social psychophysiology of vagally mediated heart rate variability: Concepts and methods in self-regulation, emotion, and interpersonal processes. Social and Personality Psychology Compass, November, 1–24. https://doi.org/10.1111/spc3.12516

Smyth, J. M. (2003). Ecological Momentary Assessment Research in Behavioral medicine. Journal of Happiness Studies, 4(1), 35–52. https://doi.org/10.1023/A:1023657221954

Smyth, N., Thorn, L., Hucklebridge, F., Evans, P., & Clow, A. (2015). Detailed time course of the cortisol awakening response in healthy participants. Psychoneuroendocrinology, 62, 200-203.

Stalder, T., Kirschbaum, C., Kudielka, B.M., Adam, E.K., Pruessner, J.C., Wüst, S., Dockray, S., Smyth, N., Evans, P., Hellhammer, D.H. and Miller, R., 2016. Assessment of the cortisol awakening response: expert consensus guidelines. Psychoneuroendocrinology, 63, pp.414-432.

Steptoe, A., & Serwinski, B. (2016). Cortisol awakening response. In Stress: Concepts, Cognition, Emotion, and Behavior (pp. 277-283).

Stone, A. A., Schwartz, J. E., Schkade, D., Schwarz, N., Krueger, A., & Kahneman, D. (2006). A population approach to the study of emotion: Diurnal rhythms of a working day examined with the day reconstruction method. Emotion, 6(1), 139–149. https://doi.org/10.1037/1528-3542.6.1.139

Stone, A. A., Schwartz, J. E., Broderick, J. E., & Shiffman, S. S. (2005). Variability of momentary pain predicts recall of weekly pain: A consequence of the peak (or salience) memory heuristic. Personality and Social Psychology Bulletin, 31(10), 1340–1346. https://doi.org/10.1177/0146167205275615

Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavorial medicine. Annals of Behavioral Medicine.

Stone, Kessler, R. C., & Haythomthwatte, J. A. (1991). Measuring Daily Events and Experiences: Decisions for the Researcher. Journal of Personality, 59(3), 575-607.

Stewart, J. C., & France, C. R. (2001). Cardiovascular recovery from stress predicts longitudinal changes in blood pressure. Biological psychology, 58(2), 105-120.

Stroud, L. R., Papandonatos, G. D., Williamson, D. E., & Dahl, R. E. (2011). Sex differences in cortisol response to corticotropin releasing hormone challenge over puberty: Pittsburgh Pediatric Neurobehavioral Studies. Psychoneuroendocrinology, 36(8), 1226-1238.

Suess, P. E., Porges, S. W., & Plude, D. J. (1994). Cardiac vagal tone and sustained attention in school-age children. Psychophysiology, 31(1), 17–22.

Sumter, S. R., Bokhorst, C. L., Miers, A. C., Van Pelt, J., & Westenberg, P. M. (2010). Age and puberty differences in stress responses during a public speaking task: Do adolescents grow more sensitive to social evaluation? Psychoneuroendocrinology, 35(10), 1510–1516. https://doi.org/10.1016/j.psyneuen.2010.05.004

Susman, E. J., Dockray, S., Granger, D. A., Blades, K. T., Randazzo, W., Heaton, J. A., & Dorn, L. D. (2010). Cortisol and alpha amylase reactivity and timing of puberty: vulnerabilities for antisocial behaviour in young adolescents. Psychoneuroendocrinology, 35(4), 557-569.

Susman, E. J., Dockray, S., Schiefelbein, V. L., Herwehe, S., Heaton, J. A., & Dorn, L. D. (2007). Morningness/eveningness, morning-to-afternoon cortisol ratio, and antisocialbehavior problems during puberty. Dev Psychol, 43(4), 811-822.

Takken, T., Stephens, S., Balemans, A., Tremblay, M. S., Esliger, D. W., Schneiderman, J., Biggar, D., Longmuir, P., Wright, V., McCrindle, B., Hendricks, M., Abad, A., Van Der Net, J., & Feldman, B. M. (2010). Validation of the Actiheart activity monitor for measurement of

activity energy expenditure in children and adolescents with chronic disease. European Journal of Clinical Nutrition, 64(12), 1494–1500. https://doi.org/10.1038/ejcn.2010.196

Tarvainen, M. P., Niskanen, J. P., Lipponen, J. A., Ranta-aho, P. O., & Karjalainen, P. A. (2014). Kubios HRV - Heart rate variability analysis software. Computer Methods and Programs in Biomedicine, 113(1). https://doi.org/10.1016/j.cmpb.2013.07.024

Taylor, J. A., Carr, D. L., Myers, C. W., & Eckberg, D. L. (1998). Mechanisms underlying very-low-frequency RR-interval oscillations in humans. Circulation, 98(6), 547–555. https://doi.org/10.1161/01.CIR.98.6.547

Thayer, J. F., Åhs, F., Fredrikson, M., Sollers, J. J., & Wager, T. D. (2012). A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health. Neuroscience and Biobehavioral Reviews, 36(2), 747–756. https://doi.org/10.1016/j.neubiorev.2011.11.009

Thayer, J. F., Åhs, F., Fredrikson, M., Sollers, J. J., & Wager, T. D. (2012b). Neuroscience and Biobehavioral Reviews A meta-analysis of heart rate variability and neuroimaging studies : Implications for heart rate variability as a marker of stress and health. Neuroscience and Biobehavioral Reviews, 36(2), 747–756. https://doi.org/10.1016/j.neubiorev.2011.11.009

Thayer, J. F., & Fischer, J. E. (2013). Heart rate variability, overnight urinary norepinephrine, and plasma cholesterol in apparently healthy human adults. International Journal of Cardiology, 162(3), 240–244. https://doi.org/https://doi.org/10.1016/j.ijcard.2011.05.058

Thayer, J. F., Hansen, A. L., Saus-Rose, E., & Johnsen, B. H. (2009). Heart rate variability, prefrontal neural function, and cognitive performance: The neurovisceral integration perspective on self-regulation, adaptation, and health. Annals of Behavioral Medicine, 37(2), 141–153. https://doi.org/10.1007/s12160-009-9101

Thayer, J. F., & Lane, R. D. (2000). A model of neurovisceral integration in emotion regulation and dysregulation. Journal of Affective Disorders, 61(3), 201–216. https://doi.org/10.1016/S0165-0327(00)00338-4

Thayer, J. F., & Sternberg, E. (2006). Beyond heart rate variability: Vagal regulation of allostatic systems. Annals of the New York Academy of Sciences, 1088, 361–372. https://doi.org/10.1196/annals.1366.014

Toone, R. J., Peacock, O. J., Smith, A. A., Thompson, D., Drawer, S., Cook, C., & Stokes, K. A. (2013). Measurement of steroid hormones in saliva: Effects of sample storage condition. Scand J Clin Lab Invest, 73(8), 615-621.

Tordjman, S., Anderson, G. M., Kermarrec, S., Bonnot, O., Geoffray, M. M., Brailly-Tabard, S., ... & Coulon, N. (2014). Altered circadian patterns of salivary cortisol in low-functioning children and adolescents with autism. Psychoneuroendocrinology, 50, 227-245.

Tran, J., Tran, R., & White, J. R. (2012). Smartphone-based glucose monitors and applications in the management of diabetes: an overview of 10 salient “apps” and a novel smartphone-connected blood glucose monitor. Clinical Diabetes, 30(4), 173- 178.

Trull, T. J., & Ebner-priemer, U. W. (2009). Research : Introduction to the Special Section. Psychological Assessment, 21(4), 457–462. https://doi.org/10.1037/a0017653.

Trull, T. J., & Ebner-Priemer, U. W. (2009). Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: introduction to the special section.

Umetani, K., Singer, D. H., McCraty, R., & Atkinson, M. (1998). Twenty-four hour time domain heart rate variability and heart rate: Relations to age and gender over nine decades. Journal of the American College of Cardiology, 31(3), 593–601. https://doi.org/10.1016/S0735-1097(97)00554-8

Valentino, K., De Alba, A., Hibel, L. C., Fondren, K., & McDonnell, C. G. (2017). Adherence to Diurnal Cortisol Sampling Among Mother-Child Dyads From Maltreating and Nonmaltreating Families. Child Maltreat, 22(4), 286-294.

Walz, L. C., Nauta, M. H., & aan het Rot, M. (2014). Experience sampling and ecological momentary assessment for studying the daily lives of patients with anxiety disorders: A systematic review. Journal of Anxiety Disorders, 28(8), 925–937. https://doi.org/10.1016/j.janxdis.2014.09.022

Wen, C. K. F., Schneider, S., Stone, A. A., & Spruijt-Metz, D. (2017). Compliance With Mobile Ecological Momentary Assessment Protocols in Children and Adolescents: A Systematic Review and Meta-Analysis. Journal of Medical Internet Research, 19(4), e132. https://doi.org/10.2196/jmir.6641

Wenze, S. J., & Miller, I. W. (2010). Use of ecological momentary assessment in mood disorders research. Clinical Psychology Review, 30(6), 794–804. https://doi.org/10.1016/j.cpr.2010.06.007

Wetherell, M. A., & Carter, K. (2014). The multitasking framework: The effects of increasing workload on acute psychobiological stress reactivity. Stress and Health, 30(2), 103–109. https://doi.org/10.1002/smi.2496

Wetherell, M. A., Craw, O., Smith, K., & Smith, M. A. (2017). Psychobiological responses to critically evaluated multitasking. Neurobiology of Stress, 7, 68–73. https://doi.org/10.1016/j.ynstr.2017.05.002

Wilde, C., Out, D., Johnson, S., & Granger, D. A. (2013). Sample collection, including participant preparation and sample handling. In The Immunoassay Handbook (Fourth Edition) (pp. 427 -440)

Wilhelm, F. H., & Grossman, P. (2010). Emotions beyond the laboratory: Theoretical fundaments, study design, and analytic strategies for advanced ambulatory assessment. Biological Psychology, 84(3), 552–569. https://doi.org/10.1016/j.biopsycho.2010.01.017

Williams, S. R., & Woodruff-Borden, J. (2015). Parent Emotion Socialization Practices and Child Self- regulation as Predictors of Child Anxiety: The Mediating Role of Cardiac Variability. Child Psychiatry And Human Development, 46(4), 512–522. http://doi.org/10.1007/s10578-014-0492-0

Yang, C. C., & Hsu, Y. L. (2010). A review of accelerometry-based wearable motion detectors for physical activity monitoring. Sensors, 10(8), 7772-7788.

Yardi, S., & Bruckman, A. (2012). Income, race, and class: exploring socioeconomic differences in family technology use. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, Texas, USA.

Young, E. A., Abelson, J. L., & Cameron, O. G. (2004). Effect of comorbid anxiety disorders on the hypothalamic-pituitary-adrenal axis response to a social stressor in major depression. Biological Psychiatry, 56(2), 113–120. https://doi.org/10.1016/j.biopsych.2004.03.017

Young, E. A., Lopez, J. F., Murphy-Weinberg, V., Watson, S. J., & Akil, H. (2000). Hormonal evidence for altered responsiveness to social stress in major depression. Neuropsychopharmacology, 23(4), 411–418. https://doi.org/10.1016/S0893-133X(00)00129-9

Zajchowski, C. A. B., Schwab, K. A., & Dustin, D. L. (2017). The Experiencing Self and the Remembering Self: Implications for Leisure Science. Leisure Sciences, 39(6), 561–568. https://doi.org/10.1080/01490400.2016.1209140

Zeiders, K. H., Doane, L. D., & Adam, E. K. (2011). Reciprocal relations between objectively measured sleep patterns and diurnal cortisol rhythms in late adolescence. J

Appendix (i) -Participant instruction pack

Participant number_____

Date _____

Daily diary web link

This link has been emailed to you also

https://goo.gl/forms/hkWhfzQR9N5f58fd2

Researcher contact information

Owen Jump
School of Applied Psychology
University College Cork

Email- owen.jump@ucc.ie

Hello

Thank you for taking the time to participate in this study. We are concerned with examining the science behind using heart rate monitoring in for mental health applications. Your participation will make a valuable contribution to our understanding of this area of science. We are grateful for the time you have taken to participate.

The following pack is intended to allow you to complete the current study. It contains information on the heart rate monitor you will be wearing as well as the daily diary measures.

If you have any questions that are not answered within the information provided or anything arises during your time wearing the monitor please contact Owen anytime at the email address below

Kind regards
Owen Jump
School of Applied Psychology
University College Cork

Email- owen.jump@ucc.ie

Daily diary

What is the daily diary?
In your pack you have been provided with a daily diary. This diary is intended to be a tool you can use to aid your recall in order to fill out the online daily diary form during the evenings you are wearing the Actiheart monitor.

This diary is **completely confidential**, no one will ever ask you to see it at any point.

How do I use the daily diary?
The best way to use the daily diary is to sit in a quiet space and think back on the previous day. As you will be asked to do it in the online form it is best to break that day down into specific episodes.

An episode can be defined as a change in location/activity/or interaction. These need to be in the order they happened over the course of the day. For example commuting to work is an episode, this may be characterised by you leaving the house and arriving at work. The diary is broken down into time periods to help you.

When should I fill out my daily diary?
The recommend time to fill out your daily diary is just before bed time, although this may present difficulties for some people, so any time in the evening.

If you have and queries regarding the daily diary please do not hesitate to contact us using the contact information inside the cover page.

Further information on how to fill out the diary is contained within the diary.

Daily online form

What is the online form

Once you have completed the series of episodes in your diary that make up the previous day you can use this information to fill out the online form.

The online form is the way we record your episodes.

The link for this has been emailed to you and is on the end of this page.

Each new episode each day will require you to click on the link and fill out the information for that episode. When you are finished that days episodes you can just close the link.

How many episodes must I fill in for each day?
The number of episodes will depend entirely on your day. There can be as many or as few as you feel accurately represents the day in question

Who will see my information on the daily diary?
The data collected from your diary will be viewed by the researcher.

However the information is anonymised using the participant number on the cover page of this document so even the researchers will not know who the information belongs to.

If you have any questions or difficulty with the online form please feel free to contact Owen

Daily diary web link

XXXXX- Link now removed

Heart rate monitor

What is the heart rate monitor?
The heart rate monitor you have been asked to wear is an Actiheart heart rate monitor. It is designed to be used in ambulatory, real life settings and to be as non- invasive as possible.

The monitor attaches to your skin on your chest and will remain there for the full 4 day duration of the study.
The heart rate monitor constantly records your heart beat and your movement.

You can take it off if you wish to do so at any time but we recommend that you NOT do this unless completely necessary. Consistent collection of data is important. If it is necessary to take it off at any time please take a note of the time you did so.

What activities can I do/not do while wearing the monitor?
We are interested in examining people in their real lives so we encourage you can do any activity you would have any planned anyway
You can wear the monitor while you sleep and it is water proof.
You can also wear the monitor during any activity such as the gym, walking etc…

What if the monitor comes loose or falls off?
Sometimes the patches we attach in the lab may come loose. If this occurs then please reattach one of the patches we have provided you with in the same position.

What do I do if the monitor becomes uncomfortable to wear?
In the unlikely event that skin irritation or discomfort occurs while wearing the monitor please feel free to remove it.

You can also take the monitor off if you wish to do so at any time but we recommend that you not do this unless COMPLETELY necessary. Consistent collection of data is important. If it is necessary to take it off at anytime please take a note of the time you did so.

What do I do with the monitor once the time is up?
Once you have completed the four day collection period we will arrange to collect the monitor from you.

If you have any further questions regarding the monitor please contact Owen

Appendix (ii) -Ethics Application

University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

ETHICS APPLICATION FORM

School of Applied Psychology UCC

(adapted from UCC Social Research Ethics Committee documentation)

Introduction

UCC academic staff and postgraduate research students who are seeking ethical approval should use this application form.

Application Checklist

This checklist includes all of the items that are required for an application to be deemed complete. In the event that any of these are not present, the application will be returned to the applicant without having been sent to review. Please ensure that your application includes all of these prior to submission. Thank you.

APPLICANT DETAILS

Name of applicant(s)	Owen Jump	**Date**	
Department/School/Unit, & Supervisor's Name	Samantha Dockray	**Phone**	
Correspondence Address	School of Applied Psycholgy UCC	**Email**	
Title of Project			

APPLICANTION DETAILS

		YES	NO
1	Do you consider that this project has significant ethical implications?		x
2	Will you describe the main research procedures to participants in advance, so that they are informed about what to expect?	x	
3	Will participation be voluntary?	x	
4	Will you obtain informed consent in writing from participants?	x	
5	Will you tell participants that they may withdraw from the research at any time and for any reason, and (where relevant) omit questionnaire items to which they do not wish to respond?	x	
6	Will data be treated with full confidentiality / anonymity (as appropriate)?	x	
7	Will data be securely held for a minimum period of seven years after the completion of a research project, in line with the University's Code of Research Conduct?	x	
8	If results are published, will anonymity be maintained and participants not identified?	x	
9	Will you debrief participants at the end of their participation (i.e. give them a brief explanation of the study)?	x	
10	Will your project involve deliberately misleading participants in any way?		x
11	Will your participants include children (under 18 years of age)?		x
12	Will your participants include people with learning or communication difficulties?		x
13	Will your participants include patients?		x
14	Will your participants include people in custody?		x
15	Will your participants include people engaged in illegal activities (e.g. drug taking; illegal Internet behaviour)?	x	
16	Is there a realistic risk of participants experiencing either physical or psychological distress?	x	
17	If yes to 16, has a proposed procedure, including the name of a contact person, been given? (see no 25)		
18	If yes to 11, is your research informed by the UCC Child Protection Policy? http://www.ucc.ie/en/ocla/policy/		

DESCRIPTION OF THE PROJECT

19. Aims of the project (briefly)

A widely used measure in stress reactivity is cardiovascular activity, specifically heart rate reactivity and heart rate variability. (Phillips & Hughes, 2011). Both have been associated with stress and health outcomes and used in studies that employ specific stressors and ambulatory settings (DePrince & Reisman, 1997). The reactivity hypothesis states that heightened cardiovascular reactivity, which is primarily an increase in heart rate and blood pressure observed in response to stress results in changes of the structure and functioning of the heart which in turn promotes sustained hypertensions. Cardiovascular reactivity to psychological stress, if prolonged or exaggerated, can promote the development of CVD (Obrist, 1981). Heightened reactivity has been linked to a number of cardiovascular outcomes, including atherosclerosis, myocardial infarction and coronary heart disease mortality (Allen et al., 1997; Everson et al.,1997). Over the past three decades the impact of several risk factors and stress buffers on reactivity have been investigated. Much research has demonstrated how personality types and traits, psychological contingencies, social support interactions, stressor tasks and task perception are associated with a particular pattern of cardiovascular reactivity that give insight into the mechanisms by which response to stress influence disease aetiology and/or progression. The paradigms for studying the impact of psychological stress on cardiovascular reactivity are inherently complex given their reliance on biological, psychological and social factors. As such, many issues remain unclear. Researchers continue to investigate the correlates of cardiovascular reactivity in order to inform theories of disease and psychological processes. The aim of this study is to examine the relationships between laboratory based measures of stress reactivity and ambulatory measures using day reconstruction methods.

20. Brief description and justification of methods and measures to be used.

This research study will use biological and psychosocial analytic methods to test heart rate variability, stress reactivity and ambulatory heart rate in a within subjects' sample from the general population.

The aim of this study is to examine the relationships between laboratory-based measures of stress reactivity and ambulatory measures using day reconstruction methods.

Full list of measures and questionnaire in Appendix (i)

If applicable, please attach research questions / copy of questionnaire / interview protocol / discussion guide / etc. materials which the Ethics Committee needs to examine in order to evaluate your application.

21. Participants: recruitment methods, number, age, gender, exclusion/inclusion criteria, detail permissions

Participants (n=50, Aged18+, males and females) have been chosen to test for associations with self-report, computer-based measures, laboratory and ambulatory settings. Participants from the general population have been chosen to examine the association between laboratory and ambulatory measures.

Participants will be accessed through social media and all exchange email system. Convenience sampling will be used. Consent will be sought in accordance with the School of Applied Psychology and Psychological Society of Ireland informed consent procedures. Participant anonymity will be paramount, and all data will be treated in accordance with data protection. There will be no inclusion or exclusion criteria from the initial survey phase. Participants will be selected for the laboratory phase based on the survey responses. There will be no exclusion criteria at this phase either. The secondary phase will examine participants based on responses to the questions in the initial survey. Participants will be assigned to the variety of groups such as Employed/Unemployed, High/Low Alcohol/Drug use, Low/High Ideation.

22. Concise statement of ethical issues raised by the project and how you intend to deal with them

The lab-based phase of the study employs the use of a computer based imaging stress task. This task elicits a moderate stress response from participants. The inclusion of a stress response in the study is key as measuring physiological reactivity in the controlled setting of a lab provides a a consistent bases of comparison for the self-report measures. It also serves as a reference point and control for the ambulatory measurement period which follows.

No other ethical issues are foreseen.

23. Arrangements for informing participants about the nature of the study (cf. Question 3 above).

Participants will be informed at the initial survey phase that the study is examining heart rate and health outcomes. Participants selected for the laboratory phase will be receive an additional contact and will be informed of their selection. Participants will not be told of any

reason for selection at this point. Participants will be told again that the study is measuring heart rate and health outcomes. After the completion of the ambulatory phase participants will be told the study was examining their laboratory and ambulatory reactivity. It is necessary not to inform participants of this until completions of the ambulatory phase as prior knowledge might introduce bias and affect behaviour in the ambulatory setting, undermining the primary aim.

Information letter and correspondence are available in appendix (ii)

If applicable, please attach the information letter / online statement / other correspondence you wish to use to inform participants about your study.

24. How you will obtain Informed Consent (cf. Question 4 above).

Informed consent will be sought at each phase of the study.
At the survey phase, informed consent, in line with online protocols will be sought. Participants will be informed as to the nature of the study (above) and right to withdraw will be fully explained. Participants will have the right to withdraw at any time. Participants will be informed of the data protection protocols in place. Participants will be informed that anonymised data may be used for publication at a later stage. Participants will have the right to have their data withdrawn for up to 6 weeks after completion of the survey.
At the laboratory phase consent will be sought again. The data collection methods, using Actiheart during the stress task and ambulatory phase will be explained. Participants will be informed of their right to cease data collection at any time without consequence. Participants will be informed of the use of anonymised data collected at this phase in subsequent publications. Participants will be informed of the right to withdraw any data for up to six weeks after collection.
If applicable, please attach (in APPENDIX 3 below) the consent form you wish to use.

25. Outline of debriefing process (cf. Question 9). If you answered YES to Question 16, give details here. State what you will advise participants to do if they should experience problems (e.g. who to contact for help).

The current study will use a Laboratory based stress task. The inclusion of a controlled Laboratory based stress task is required to serve as a comparison/ baseline to the ambulatory phase. The Montreal Imaging Stress Task is a computer based task designed to elicit a moderate stress response to a math based task. The task is based on the numerical section of the trier social stress task but provides more uniform stimulus to participants. There is a small risk of participant distress due t exposure to this task. In the event of distress, the task will cease immediately. All participants will be fully de-briefed after the task as to the nature of the task. Participants will be informed that they were participating in a stress task. De brief procedures inline with the the de brief script for the Trier Social Stress task (Birkett, 2011) will be followed. This procedure has been designed to negate any possible negative consequences of the stress task.

26. Estimated start date and duration of project

Signed _____ Date _____
Applicant

Signed _____ Date _____
Research Supervisor/Principal Investigator (if applicable)

Please submit this form and attachments to adminapsych@ucc.ie, with the words ethics application (followed by your full name) in the subject line). Please include a scan of the signatures required. No hard copies are required.

This form is adapted from pp. 13-14 of Guidelines for Minimum Standards of Ethical Approval in Psychological Research (British Psychological Society, July, 2004)

Last update: September 2015

Demographics

Please fill out all of the following information.

1. Age

2. Gender

3. Are you currently employed? If not how long have you been unemployed?

4. Are you currently looking for employment?

5. Out of the last ten years how many have you been in steady employment?

6. What is your primary area of training/skill/education

An Ultra-Brief Screening Scale for Anxiety and Depression: the PHQ-4

(Kroenke, Spitzer, Williams, & Löwe., 2009).

Over the past two weeks have you been bothered by these problems?

	Not at all	Several days	More days than not	Nearly every day
Feeling nervous, anxious or on edge	0	1	2	3
Not being able to stop or control worrying	0	1	2	3
Feeling down, depressed or hopeless	0	1	2	3
Little interest or pleasure in doing things	0	1	2	3

Simple Screening Instrument for Substance Abuse (Center Substance Abuse Treatment, 1994)

Directions: The questions that follow are about your use of alcohol and other drugs. Your answers will be kept private. Mark the response that best fits for you. Answer the questions in terms of your experiences in the past 6 months.
During the last 6 months…
1. Have you used alcohol or other drugs? (Such as wine, beer, hard liquor, pot, coke, heroin or other opioids, uppers, downers, hallucinogens, or inhalants)
___ Yes ___ No
2. Have you felt that you use too much alcohol or other drugs?
___ Yes ___ No
3. Have you tried to cut down or quit drinking or using alcohol or other drugs?
___ Yes ___ No
4. Have you gone to anyone for help because of your drinking or drug use? (Such as Alcoholics Anonymous, Narcotics Anonymous, Cocaine Anonymous, counselors, or a treatment program.)
___ Yes ___ No
5. Have you had any health problems? For example, have you:
___ Had blackouts or other periods of memory loss?
___ Injured your head after drinking or using drugs?
___ Had convulsions, delirium tremens (“DTs”)?
___ Had hepatitis or other liver problems?

___ Felt sick, shaky, or depressed when you stopped?
___ Felt “coke bugs” or a crawling feeling under the skin after you stopped using drugs?
___ Been injured after drinking or using?
___ Used needles to shoot drugs?
6. Has drinking or other drug use caused problems between you and your family or friends?
___ Yes ___ No
7. Has your drinking or other drug use caused problems at school or at work?
___ Yes ___ No
8. Have you been arrested or had other legal problems? (Such as bouncing bad checks, driving while intoxicated, theft, or drug possession.)
___ Yes ___ No
9. Have you lost your temper or gotten into arguments or fights while drinking or using other drugs?
___ Yes ___ No
10. Are you needing to drink or use drugs more and more to get the effect you want?
___ Yes ___ No
11. Do you spend a lot of time thinking about or trying to get alcohol or other drugs?
___ Yes ___ No
12. When drinking or using drugs, are you more likely to do something you wouldn't normally do, such as break rules, break the law, sell things that are important to you, or have unprotected sex with someone?
___ Yes ___ No
13. Do you feel bad or guilty about your drinking or drug use?
___ Yes ___ No
The next questions are about your lifetime experiences.
14. Have you ever had a drinking or other drug problem?
___ Yes ___ No
15. Have any of your family members ever had a drinking or drug problem?
___ Yes ___ No
16. Do you feel that you have a drinking or drug problem now?
___ Yes ___ No
Thanks for filling out this questionnaire.

Day reconstruction method (DRM). (Kahneman, Krueger, Schkade, Schwarz, & Stone, 2004) The Day Reconstruction Method is a used for assessing experiences and affect in daily life.. The technique uses retrospective reconstruction using a diary of set time periods to assess activities and affect of the previous day (Stone & Schwartz, 2006). This measure has been included as it represents a less invasive method of daily experience. It has been validated against other momentary assessment and show to be a reliable method of for monitoring affect and experience over defined epochs (Dockray, Grant , Stone , Kahneman, Wardle , & Steptoe, 2010).

Actiheart (Brage, Brage, Franks, Ekelund, & Wareham, 2005) Is a validated non-invasive method of measuring physical activity and energy expenditure (PAEE). It is attached over the skin and can be used for long periods of activity in an ambulatory setting. Actiheart is used to measure both movement and EKG heart rate. The result is an overview of the participant movement and heart rate during a specified period. Information is downloaded to accompanying computer software to allow examination of heart rate (Stone & Schwartz, 2006)

Momentary Affect Scale
Please rate on the scale below the number that best describes how you currently feel

	Not at all						Very Much
Relaxed	0.	1.	2.	3.	4.	5.	6
Stressed	0.	1.	2.	3.	4.	5.	6
In control	0.	1.	2.	3.	4.	5.	6
Anxious	0.	1.	2.	3.	4.	5.	6

Information letter

University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

School of Applied Psychology
University College Cork
Phone: 0876569209
E-Mail: owen.jump@ucc.ie
Web::http://www.ucc.ie/en/apsych/

Dear Participant

My name is Owen Jump and I am a PhD candidate in the School of Applied Psychology in University College Cork. Our team is conducting a study that is interested in understanding the use of heart rate in mental health applications

The use of heart rate monitoring has become widespread in sports applications and more recently in helping people monitor stress. As such we are interested in studying these effects in peoples' daily lives. That is why we require your help in participating in this study.

Initially you will just be required to fill out a short survey (it takes less than 10 minutes to complete). We will then be randomly selecting participants from the survey to wear heart rate monitors for one week in their daily lives.

In the study we will be using a number of measures. These measures are; Actiheart monitors that measure heart rate and some short surveys.

If you are interested in participating, please fill out the survey and I will get in contact with you soon after.

Sincerely,
Owen Jump
School of Applied Psychology
University College Cork

INFORMATION SHEET

University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Purpose of the Study. As part of the requirements for my PhD at UCC, I have to carry out a research study. The study is concerned with examining the use of heart rate measures in mental health applications.

What will the study involve? The study will involve two stages. The first is to fill out the short questionnaire attached to this email. This survey is very short and takes between 5-10 minutes. The second phase will be study requiring you to attend the School of Applied Psychology and wear a heart rate monitor. You will then be asked to wear the heart rate monitor for up to one week and fill out a daily diary.

Agreeing to phase 1 does not mean you automatically agree to phase 2. I will contact you using the information you provide in the survey and discuss further participation.

Why have you been asked to take part? You have been asked to take part to help in this study because we are seeking healthy participants from the general population.

Do you have to take part? Participation is completely voluntary. You may withdraw at any time without any consequence. You may also have your data removed at any time of your choosing.

Will your participation in the study be kept confidential? Your participation and all subsequent data will be confidential. Although we collect contact information it is only to allow us to keep records. Data protection guidelines which outline how data is treated will ensure

confidentialty. Some anonymized data and numerical data will may be used for publication at a later stage.

Who has reviewed this study? This study has been reviewed by University College Cork School of Applied Psychology Ethics Board. All ethical issues as set out by the School and the Psychological society of Ireland code of ethics will be adhered to.

Any further queries? If you need any further information, you can contact me @ owen.jump@ucc.ie

If you agree to take part in the study, please sign the consent form overleaf.

CONSENT FORM

University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

This consent form is designed with qualitative research in mind. Where quantitative methods are used, issues such as quotations and audio-recording do not arise.

I……agree to participate in Owen Jumps' research study.

The purpose and nature of the study has been explained to me in writing.

I am participating voluntarily.
I give permission for my heart rate data to be used for publication/dissemination at a later time
I agree to take part in Computer based laboratory tasks
I agree to wear the Actiheart Monitor for up to one week

I understand that I can withdraw from the study, without repercussions, at any time, whether before it starts or while I am participating.

I understand that I can withdraw permission to use the data after six weeks , in which case the material will be deleted.

I understand that anonymity will be ensured in the write-up by disguising my identity.

I understand that disguised extracts from my interview may be quoted in the thesis and any subsequent publications if I give permission below:

(Please tick one box:)
I agree to quotation/publication of extracts from my interview ☐

I do not agree to quotation/publication of extracts from my interview ☐

Signed: ……. Date: ……..
PRINT NAME: …….

Authorisation

Date when this research was considered by the Ethics Committee of the School of Applied Psychology

Result (please check one):

Approved	x
Approved with minor comments (resubmission is not required)	
Approved pending clarification (a list of sections and required clarifications must be made below or appended)	
Approved pending approval from external body (the body or bodies from which approval is pending must be specified)	
Not approved	
Referred to Social Research Ethics Committee (SREC)	
Other	

Feedback for the applicant:

Dear Owen

Thank you for resubmitting the above Ethics Application. It has been approved by the Ethics Committee.

We wish you all the best with your study.

Kind Regards

Mags Creedon

Mags Creedon| School of Applied Psychology | Enterprise Centre, North Mall Campus | University College Cork : +353 21 4904604|